CRM 86-228 / November 1986

RESEARCH MEMORANDUM

EVALUATING THE APPROPRIATENESS OF THE NUMERICAL OPERATIONS AND MATH KNOWLEDGE SUBTESTS IN THE AFQT

Milton H. Maier
Catherine M. Hiatt

A Division of CNA Hudson Institute
CENTER·FOR·NAVAL·ANALYSES
4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268 • (703) 824-2000

Work conducted under contract N00014-83-C-0725.

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

A180266

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) CRM 86-228	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Center for Naval Analyses 6b. OFFICE SYMBOL (If applicable) CNA	7a. NAME OF MONITORING ORGANIZATION Commandant of the Marine Corps (Code RDS)
6c. ADDRESS (City, State, and ZIP Code) 4401 Ford Avenue Alexandria, Virginia 22302-0268	7b. ADDRESS (City, State, and ZIP Code) Headquarters, Marine Corps Washington, D.C. 20380
8a. NAME OF FUNDING / ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable) ONR	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-83-C-0725
8c. ADDRESS (City, State, and ZIP Code) 800 North Quincy Street Arlington, Virginia 22217	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 65153M; PROJECT NO. C0031; TASK NO.; WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)

Evaluating the Appropriateness of the Numerical Operations and Math Knowledge Subtests in the AFQT

12. PERSONAL AUTHOR(S)
Milton H. Maier, Catherine M. Hiatt

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM TO	November 1986	94

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The Numerical Operations (NO) subtest, used in the Armed Forces Qualification Test (AFQT), has a history of causing problems for the Armed Services Vocational Aptitude Battery (ASVAB) score scale. Two adjustments to the ASVAB score scale have been required since NO was made part of the AFQT in 1980. The purpose of this report is to present evidence that the NO subtest should be deleted from the AFQT and be replaced by the Math Knowledge (MK) subtest. Data are presented showing that the AFQT scores in 1984 were inflated by approximately 3 percentile score points.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Lt.Col. G. W. Russell	22b. TELEPHONE (Include Area Code) (202) 694-3491 22c. OFFICE SYMBOL RDS-40

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted.

All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

A Division of CNA Hudson Institute

CENTER·FOR·NAVAL·ANALYSES

4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268 • (703) 824-2000

4 November 1986

MEMORANDUM FOR DISTRIBUTION LIST

Subj: Center for Naval Analyses Research Memorandum 86-228

Encl: (1) CNA Research Memorandum 86-228, "Evaluating the Appropriateness of the Numerical Operations and Math Knowledge Subtests in the AFQT," by Milton H. Maier and Catherine M. Hiatt, November 1986

1. Enclosure (1) is forwarded as a matter of possible interest.

2. The Armed Forces Qualification Test (AFQT) is used by the Armed Services as a measure of general trainability. AFQT scores are derived from various subtests contained within the Armed Services Vocational Aptitude Battery (ASVAB). One of these subtests, Numerical Operations (NO), has caused repeated problems in the calibration of AFQT. This Research Memorandum examines the score inflation in operational test scores due to the NO subtest. The appropriateness of retaining NO in the AFQT is assessed.

William H. Sims
Director, Manpower and Training Program
Marine Corps Operations
Analysis Group

Distribution List:
Reverse Page

Subj: Center for Naval Analyses Research Memorandum 86-228

Distribution List

SNDL

A1	ASSTSECNAV MRA
A1	DASN - MANPOWER (2 copies)
A6	HQMC MPR
	Attn: Deputy Chief of Staff for Manpower (2 copies)
	Attn: Director, Personnel Procurement Division (2 copies)
	Attn: Director, Manpower Plans and Policy Division (2 copies)
	Attn: Director, Personnel Management Division (2 copies)
A6	HQMC TRNG (2 copies)
A6	HQMC RD&S (2 copies)
A6	HQMC RA (2 copies)
A6	HQMC AVN (2 copies)
E3D1	CNR
E3D5	NAVPERSRANDCEN
	Attn: Director, Manpower and Personnel Laboratory
	Attn: Technical Library
FF38	USNA
	Attn: Nimitz Library
FF42	NAVPGSCOL
FF44	NAVWARCOL
FJA1	COMNAVMILPERSCOM
FJB1	COMNAVCRUITCOM
FT1	CNET
V12	CG MCDEC

OPNAV

OP-01
OP-11
OP-12
OP-13

Other

Joint Service Selection and Classification Working Group (11 copies)

CRM 86-228 / November 1986

EVALUATING THE APPROPRIATENESS OF THE NUMERICAL OPERATIONS AND MATH KNOWLEDGE SUBTESTS IN THE AFQT

Milton H. Maier
Catherine M. Hiatt

Marine Corps Operations Analysis Group

A Division of CNA Hudson Institute
CENTER·FOR·NAVAL·ANALYSES
4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268

ABSTRACT

The Numerical Operations (NO) subtest, used in the Armed Forces Qualification Test (AFQT), has a history of causing problems for the Armed Services Vocational Aptitude Battery (ASVAB) score scale. Two adjustments to the ASVAB score scale have been required since NO was made part of the AFQT in 1980. The purpose of this report is to present evidence that the NO subtest should be deleted from the AFQT and be replaced by the Math Knowledge (MK) subtest. Data are presented showing that the AFQT scores in 1984 were inflated by approximately 3 percentile score points.

EXECUTIVE SUMMARY

BACKGROUND

The Numerical Operations (NO) subtest, used in the Armed Forces Qualification Test (AFQT), has a history of causing problems for the Armed Services Vocational Aptitude Battery (ASVAB) score scale. Two adjustments to the ASVAB score scale have been required since NO was made part of the AFQT in 1980. The purpose of this report is to present evidence that the NO subtest should be deleted from the AFQT and be replaced by the Math Knowledge (MK) subtest.

Design of Testing Materials

The first adjustment was required when a new score scale for the ASVAB, called the 1980 score scale, was being constructed. The ASVAB was administered in 1980 to a nationally representative sample of American youth, called the 1980 Youth Population, for the purpose of constructing a new score scale. During the construction, the NO scores of military examinees were found to be higher than the NO scores for examinees in the 1980 Youth Population, whereas on other tests the military examinees scored lower. After much searching, the reason was found to lie in the design of the answer sheets. As shown in figure I, the answer sheet for the 1980 Youth Population required examinees to fill in circles, whereas military examinees filled in rectangles. Because the time limit for the NO subtest is only 3 minutes, the additional time required to fill in the circles lowered the scores. In addition, the NO items in the test booklet were arranged in seven columns of seven items each plus one item in the eighth column, in contrast to the layout of the answer sheet used by the 1980 Youth Population, as shown in figure I. Examinees could more easily keep track of item and response locations with the military testing materials. Scores on the NO subtest had to be adjusted before the 1980 score scale was even introduced, on 1 October 1984.

The second adjustment was made on 1 July 1986. A new version of the ASVAB, forms 11, 12, and 13 (ASVAB 11/12/13), was introduced on 1 October 1984, along with the new 1980 score scale. An Initial Operational Test and Evaluation (IOT&E) was conducted immediately after the new ASVAB and score scale were introduced. The results of the IOT&E showed that ASVAB 11/12/13 was not accurately calibrated to the 1980 Youth Population; the AFQT scores were approximately 2 percentile score points too low. The adjustment on 1 July 1986 raised the AFQT scores by 2 percentile

Answer spaces for the 1980 Youth Population

PART 5 — NUMERICAL OPERATIONS

1 Ⓐ Ⓑ Ⓒ Ⓓ	11 Ⓐ Ⓑ Ⓒ Ⓓ	21 Ⓐ Ⓑ Ⓒ Ⓓ	31 Ⓐ Ⓑ Ⓒ Ⓓ	41 Ⓐ Ⓑ Ⓒ Ⓓ
2 Ⓐ Ⓑ Ⓒ Ⓓ	12 Ⓐ Ⓑ Ⓒ Ⓓ	22 Ⓐ Ⓑ Ⓒ Ⓓ	32 Ⓐ Ⓑ Ⓒ Ⓓ	42 Ⓐ Ⓑ Ⓒ Ⓓ
3 Ⓐ Ⓑ Ⓒ Ⓓ	13 Ⓐ Ⓑ Ⓒ Ⓓ	23 Ⓐ Ⓑ Ⓒ Ⓓ	33 Ⓐ Ⓑ Ⓒ Ⓓ	43 Ⓐ Ⓑ Ⓒ Ⓓ
4 Ⓐ Ⓑ Ⓒ Ⓓ	14 Ⓐ Ⓑ Ⓒ Ⓓ	24 Ⓐ Ⓑ Ⓒ Ⓓ	34 Ⓐ Ⓑ Ⓒ Ⓓ	44 Ⓐ Ⓑ Ⓒ Ⓓ
5 Ⓐ Ⓑ Ⓒ Ⓓ	15 Ⓐ Ⓑ Ⓒ Ⓓ	25 Ⓐ Ⓑ Ⓒ Ⓓ	35 Ⓐ Ⓑ Ⓒ Ⓓ	45 Ⓐ Ⓑ Ⓒ Ⓓ
6 Ⓐ Ⓑ Ⓒ Ⓓ	16 Ⓐ Ⓑ Ⓒ Ⓓ	26 Ⓐ Ⓑ Ⓒ Ⓓ	36 Ⓐ Ⓑ Ⓒ Ⓓ	46 Ⓐ Ⓑ Ⓒ Ⓓ
7 Ⓐ Ⓑ Ⓒ Ⓓ	17 Ⓐ Ⓑ Ⓒ Ⓓ	27 Ⓐ Ⓑ Ⓒ Ⓓ	37 Ⓐ Ⓑ Ⓒ Ⓓ	47 Ⓐ Ⓑ Ⓒ Ⓓ
8 Ⓐ Ⓑ Ⓒ Ⓓ	18 Ⓐ Ⓑ Ⓒ Ⓓ	28 Ⓐ Ⓑ Ⓒ Ⓓ	38 Ⓐ Ⓑ Ⓒ Ⓓ	48 Ⓐ Ⓑ Ⓒ Ⓓ
9 Ⓐ Ⓑ Ⓒ Ⓓ	19 Ⓐ Ⓑ Ⓒ Ⓓ	29 Ⓐ Ⓑ Ⓒ Ⓓ	39 Ⓐ Ⓑ Ⓒ Ⓓ	49 Ⓐ Ⓑ Ⓒ Ⓓ
10 Ⓐ Ⓑ Ⓒ Ⓓ	20 Ⓐ Ⓑ Ⓒ Ⓓ	30 Ⓐ Ⓑ Ⓒ Ⓓ	40 Ⓐ Ⓑ Ⓒ Ⓓ	50 Ⓐ Ⓑ Ⓒ Ⓓ

Answer spaces for military examinees

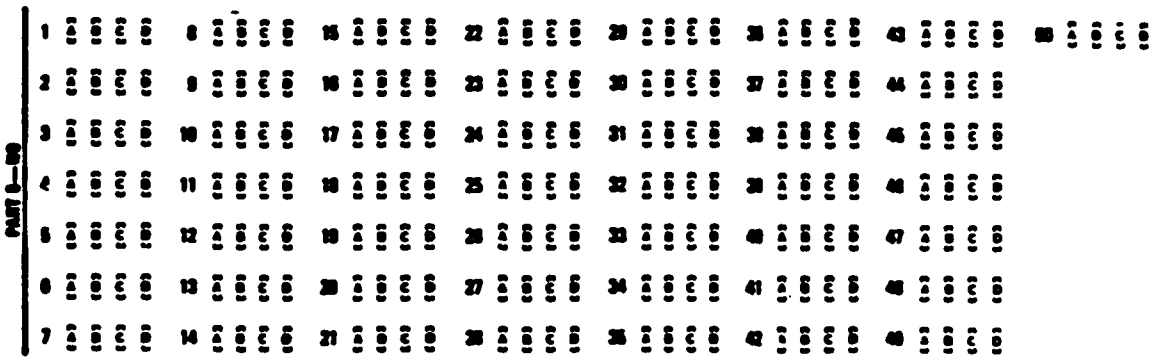


FIG. I: ANSWER SPACES FOR THE NUMERICAL OPERATIONS SUBTEST USED WITH THE 1980 YOUTH POPULATION AND MILITARY EXAMINEES

score points. The discrepancy in scores was again due to problems with the NO subtest. The type font for the test booklets was inadvertently changed between the initial calibration of ASVAB 11/12/13 and introduction of forms for operational use on 1 October 1984. The type font for the operational test booklets was more difficult to read, and the NO scores were lowered. A similar problem was also found in 1983 with forms 10X and 10Y of the ASVAB.

Until these problems with NO scores were discovered, no one suspected that speeded tests were so sensitive to differences in the design of testing materials. The design problems, in principle, can be corrected by exercising more careful quality control over the printing of the materials. There are, however, inherent defects in the NO subtest that do not lend themselves to correction.

Inherent Defects of the NO Subtest

The NO subtest, as a speeded test, is inherently defective when used in an ongoing testing program, such as testing applicants for enlistment to determine their mental qualifications. The reason is that it contains 50 test items (adding, subtracting, multiplying, and dividing two 1- or 2-digit numbers) to be completed in a 3-minute time limit, and examinees can easily improve their test scores. Strategies to improve NO scores include:

- Memorizing the answers to the arithmetic problems, say all combinations up to 12
- Keeping track of question and answer locations in the test booklet and on the answer sheet;
- Working as quickly as possible
- Continuing to answer items after time is called.

All these test-taking strategies, except the last, are legitimate practices, and do not constitute cheating in the conventional sense. Most examinees, however, do need to be taught the strategies for taking the test. The result is that their NO scores, and consequently their AFQT scores, are inflated relative to the scores of the 1980 Youth Population. The inflated scores result in classification errors that overestimate the aptitude of the examinees who use these, and perhaps other, test-taking strategies.

The first analysis in this report quantifies the amount of inflation of the operational NO and AFQT scores. The second evaluates the impact of the inflated AFQT scores and an alternative AFQT, in which the MK subtest replaces the NO subtest, on the qualification rates of applicants for enlistment. This alternative definition of the AFQT was proposed by the Joint Services Selection and Classification Working Group.

AMOUNT OF INFLATION IN THE NO AND AFQT SCORES

Inflation of the NO scores was measured by comparing score distributions of applicants for active duty in the IOT&E for ASVAB 11/12/13, conducted in October and November 1984, to those of the 1980 Youth Population. Prior to equating the NO scores, the two groups were weighted to have the same distribution of general aptitude (verbal, math, and technical). NO scores with the same cumulative frequency in the two weighted groups were set equal to each other. Separate equatings were performed for males and females. The amount of inflation for males in the 1984 IOT&E group is shown in figure II. The inflation is largest in the lower half of the NO scores (from 12 through 20). The maximum inflation is 8 points for males and 7 points for females. In contrast, the MK scores are not inflated.

To estimate the impact of this inflation on AFQT, the NO scores for people in the 1984 IOT&E were adjusted downward by the amount of the inflation. A new set of AFQT scores was then computed for each person, where AFQT is defined as the sum of the Word Knowledge, Paragraph Comprehension, Arithmetic Reasoning, and one-half of the Numerical Operations subtest scores. One set of AFQT scores includes the operational NO scores, which are inflated, and the other includes the adjusted NO scores, which have the inflation removed. The other AFQT subtest scores, of course, remain unchanged; differences in the two AFQT distributions reflect the inflated NO scores.

The amount of inflation in the AFQT scores is shown in figure III. The amount of inflation is up to 3 percentile score points for both males and females. The large differences occur in the middle of the scale, percentile scores 40 to 60.

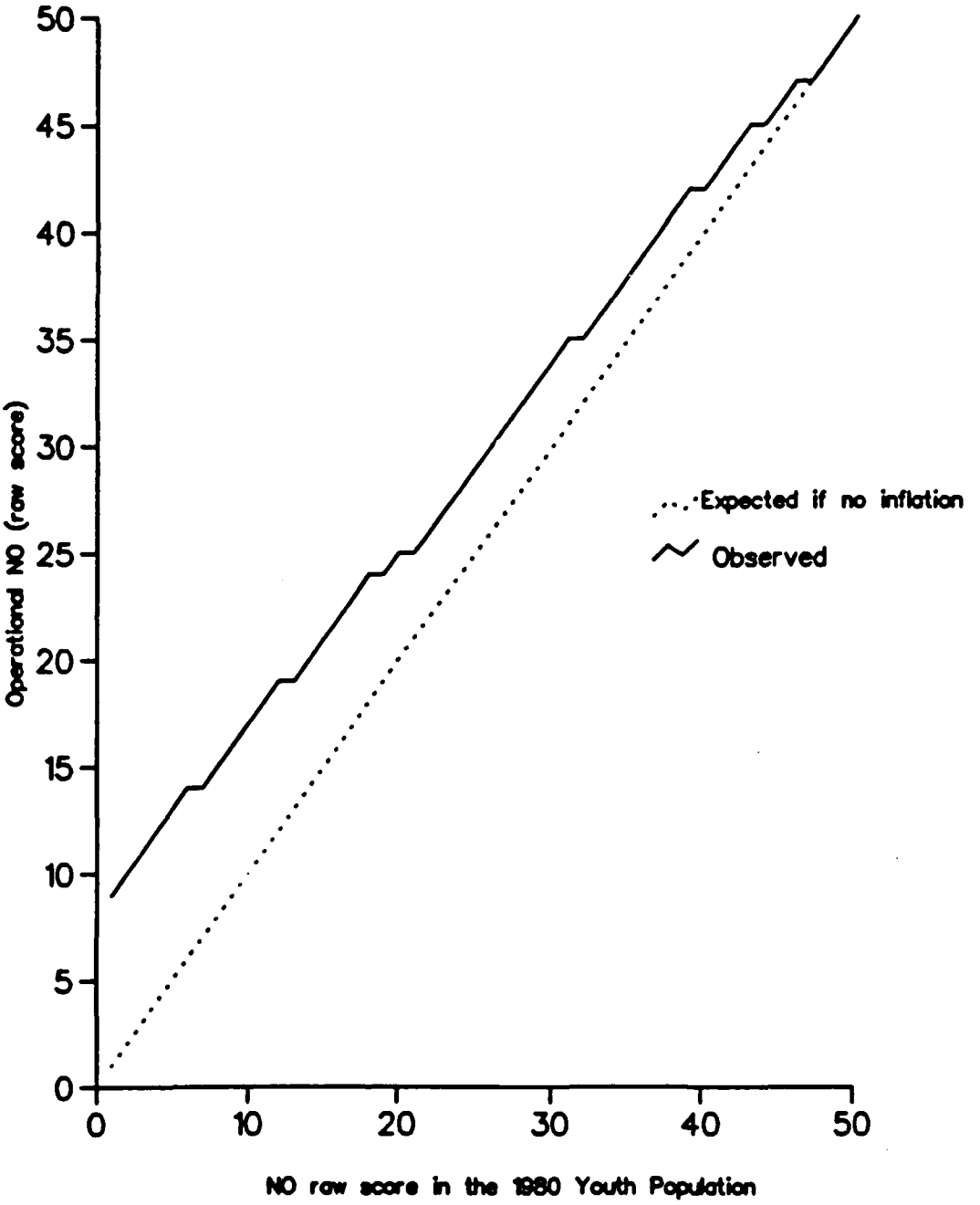


FIG. II : INFLATION IN NO SCORES FOR MALES (1984 IOT&E)

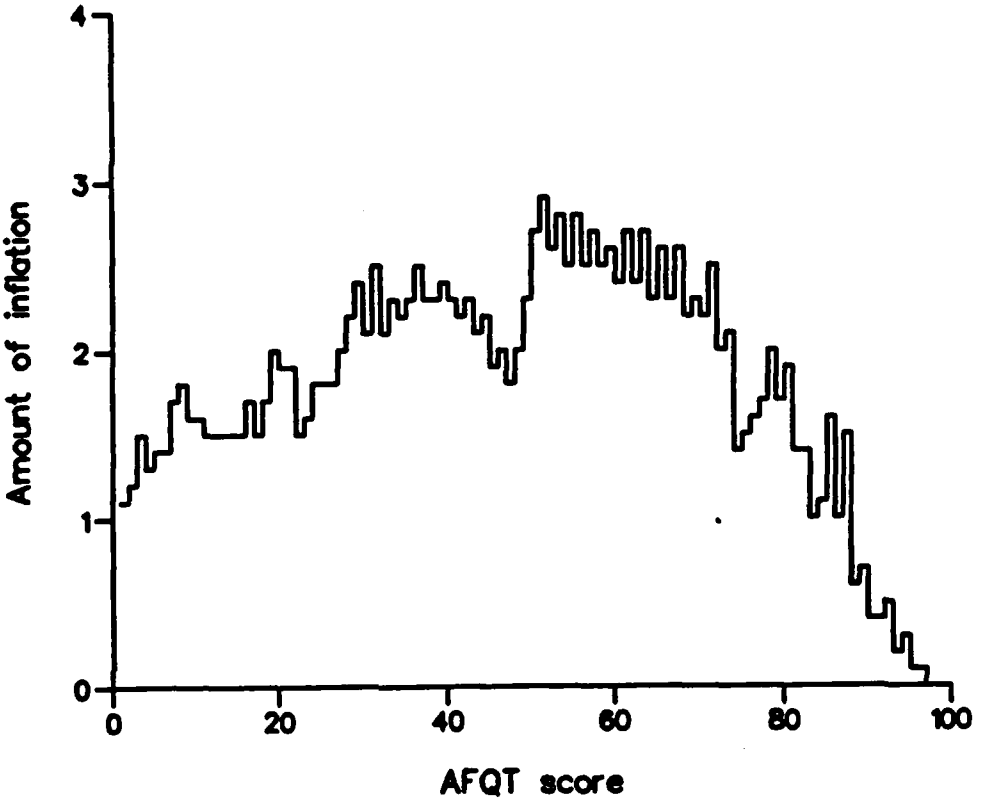


FIG. III: AMOUNT OF INFLATION IN AFQT PERCENTILE SCORES
(1984 IOT&E)

This inflation means that, on average, the AFQT scores of applicants are up to 3 percentile score points too high. The classification errors work in the direction of overestimating aptitude, and more people qualify for enlistment, bonuses, and guaranteed assignments than would be the case if NO and AFQT were accurately calibrated to the 1980 Youth Population.

COMPARISON OF QUALIFICATION RATES

Personnel managers are concerned about replacing NO with MK in the AFQT because of the impact the change may have on qualification rates. Females tend to have higher scores than males on speeded tests. Because NO is a speeded test, fewer females would qualify for enlistment and bonuses under the proposed AFQT, which contains MK, than under the current version.

Comparison of qualification rates, however, cannot be based on operational scores because the inflated NO scores magnify true differences between the current and proposed definitions of the AFQT. Consequently, qualification rates were computed for three sets of AFQT scores: current operational, with inflated NO scores; current adjusted, with NO scores adjusted for inflation; and proposed, with MK replacing NO. The qualification rates of applicants in the 1984 IOT&E were computed to evaluate the impact of the proposed AFQT.

The qualification rates for the total number of males and black males are shown in figure IV, and the rates for the total number of females and black females are shown in figure V. The percentage of qualified applicants is shown at percentile scores of 31 and 50. Applicants for all services combined were used to compute these qualification rates; the rates for each service are shown in the main text.

More males, both for the total group and blacks, could qualify on the proposed AFQT than on either the current adjusted or current operational AFQT. This means that the number of qualified males would increase if the proposed AFQT were used, even when compared with the current inflated AFQT scores.

For the total group of females, the qualification rates at a percentile score of 31 are about the same on the adjusted and proposed AFQT scores and about 2 percentage points lower than on the operational (inflated) scores. At a percentile score of 50, the qualification rate is about 1 percentage point higher

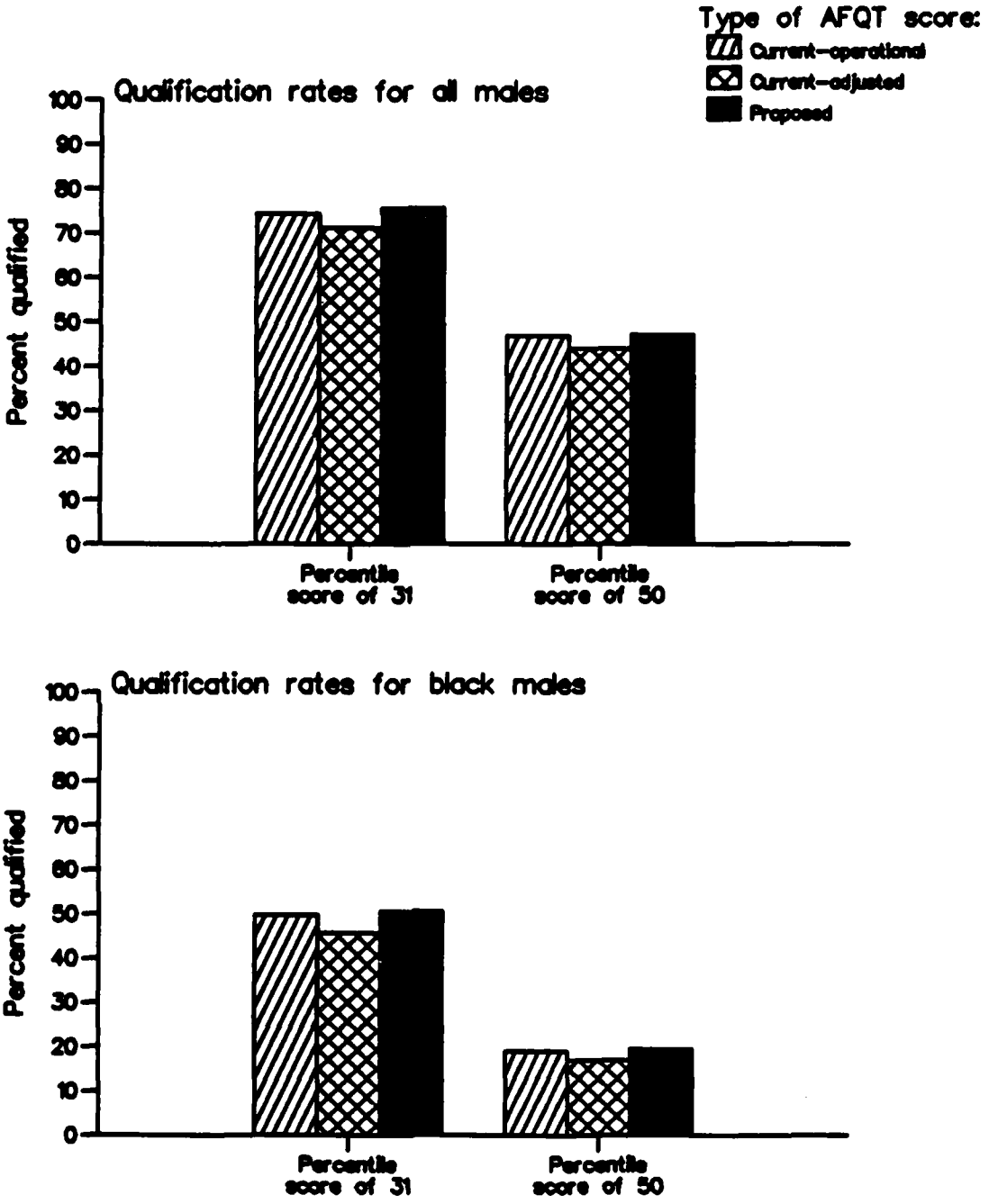


FIG IV: QUALIFICATION RATES FOR MALES
(1984 IOT&E)

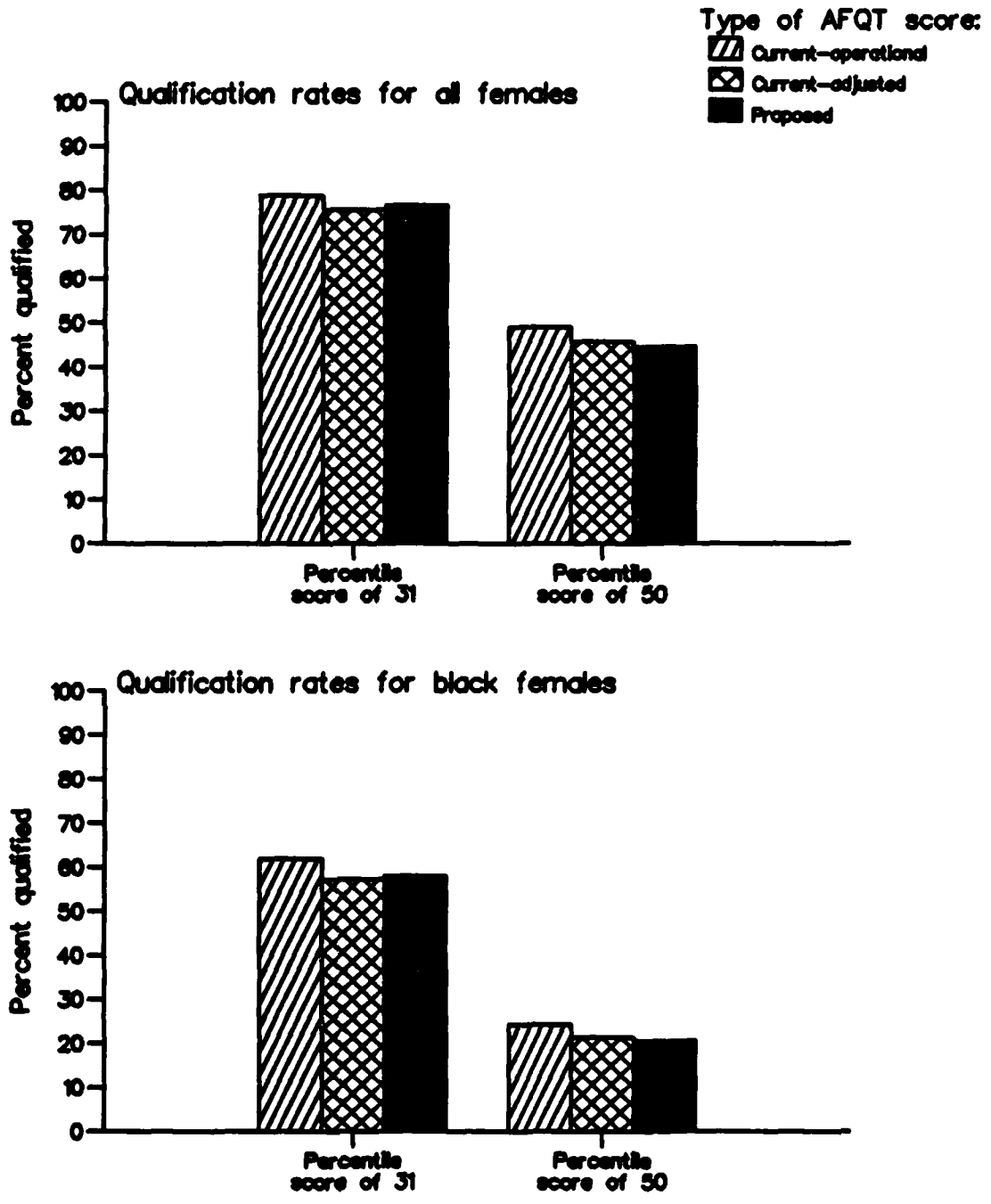


FIG V: QUALIFICATION RATES FOR FEMALES
(1984 IOT&E)

on the adjusted AFQT than on the proposed (46 versus 45 percent). At a percentile score of 50, the rates on both the adjusted and proposed AFQT are lower than on the operational AFQT (3 percentage points on the adjusted and 4 on the proposed). The pattern of qualification rates for black females is similar to that for the total group.

The differences in qualification rates among the AFQTs tend to be equal for the three racial groups—whites, blacks, and others (shown in the main text). At both percentile score levels, the qualification rates on the operational AFQT are about 3 percentage points higher than on the adjusted AFQT. This difference reflects the inflation of the current AFQT.

Because of the inflation of the operational AFQT scores (up to 3 percentile score points), the most accurate estimate of the adverse impact of the proposed AFQT on females is the difference between the adjusted and proposed AFQT scores.

CONCLUSIONS AND RECOMMENDATIONS

Conclusions from the analysis are as follows:

- Operational NO scores have, over time, become inflated relative to the 1980 Youth Population. As a result, AFQT scores in 1984 were inflated by approximately 3 percentile score points and cannot be said to be on the 1980 score scale.

- Inflation of the NO scores raises questions about their predictive validity. The NO subtest is in several aptitude composites, which are used in assigning recruits to occupational specialties. The predictive validity of the inflated NO scores needs to be evaluated.

- The NO subtest is inherently flawed for use in an operational testing program where examinees have prior information about the types of test content. Examinees can legitimately prepare themselves to improve their speeded test scores without improving the underlying ability that would also improve performance in their occupational specialties. Continued use of NO in the AFQT would negate the efforts to maintain the accuracy of the ASVAB score scale.

- Adjusting the inflated operational AFQT scores is not practical because the adjustments would be accurate only on average, but not necessarily for any individual. To the extent that individuals differ in how much they use the test-taking strategies, their NO and AFQT scores are inflated by differing amounts. Thus any adjustment would be unfair to some individuals.

- Changing the AFQT by replacing NO with MK would result in an apparent loss of qualified recruits only because of the people who had qualifying AFQT scores based on the inflated NO scores. Once the distortion due to inflation is removed, use of the proposed AFQT would actually increase the percentage of qualified males and females at the percentile score of 31.

Recommendations from the analysis are the following:

- Delete NO from the AFQT and replace it with MK.

- Evaluate the predictive validity of the inflated NO scores to determine whether this subtest should be retained in the ASVAB.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Continued)

LIST OF TABLES

PURPOSE

The purpose of this report is to present evidence that supports the recommendation of the Joint Services Selection and Classification Working Group[1] that the Numerical Operations (NO) subtest should be deleted from the Armed Forces Qualification Test (AFQT) and replaced by the Math Knowledge (MK) subtest. The recommendation is controversial among personnel managers in the military services because of its possible adverse impact on the qualification rates of females and racial minorities. The NO subtest, however, has a history of problems arising from the design and format of answer sheets and test booklets; these problems are reviewed in this report. In addition, the NO subtest suffers from inherent defects that inflate the scores and introduce errors into the personnel decisions that are based in part on the NO scores. The bulk of this report is devoted to documenting the magnitude of these inherent defects and evaluating their impact on the qualification rates of applicants for enlistment.

PROBLEMS ARISING FROM THE DESIGN AND FORMAT OF TESTING MATERIALS

The NO subtest was made a part of the AFQT in October 1980, when forms 8, 9, and 10 of the Armed Services Vocational Aptitude Battery (ASVAB 8/9/10) were introduced. The subtests that compose the ASVAB are shown in table 1. They are arranged by type of content (verbal, math, technical, and speeded). The number of items and time limits are also shown. All subtests except the speeded ones have generous time limits. The NO subtest had been part of the earlier version of the ASVAB, but it had not been included in the AFQT. Because the AFQT score is used as the first screen for enlistment in the armed services, recruiters and applicants for enlistment tend to focus their efforts to improve performance on the subtests that compose the AFQT. These are Word Knowledge (WK), a 35-item vocabulary test, Paragraph Comprehension (PC), a 15-item reading comprehension test, Arithmetic Reasoning (AR), a 35-item test of word problems, and NO, a speeded test of 50 simple arithmetic problems, with a time limit of 3 minutes.

1. The Working Group is composed of technical and policy representation from each service and the Office of the Secretary of Defense. It is responsible for the development and operation of the Armed Services Vocational Aptitude Battery and other instruments used in the selection and classification of recruits. The recommendation was not supported by the Army representatives to the Working Group.

The NO items involve addition, subtraction, multiplication, and division of two 1- or 2-digit numbers.

TABLE 1

CONTENT OF THE ASVAB

Subtest	Number of items	Time limit (min)
Verbal		
General Science	25	11
Word Knowledge	35	11
Paragraph Comprehension	15	13
Math		
Arithmetic Reasoning	30	36
Math Knowledge	25	24
Technical		
Auto/Shop Information	25	11
Mechanical Comprehension	25	19
Electronics Information	20	9
Speed		
Numerical Operations	50	3
Coding Speed	84	7

Problems with the NO subtest emerged shortly after ASVAB 8/9/10 was introduced. The first indication that NO scores were affected by the test materials was discovered in 1983, when a new score scale for the ASVAB was being constructed [1]. Form 8 of the ASVAB was administered in the summer and fall of 1980 to a nationally representative sample of American youth. A new ASVAB score scale, called the 1980 score scale or metric, was constructed on this population [2]. A comparison of people tested with the military testing materials and the 1980 Youth Population showed that the military examinees scored lower on all ASVAB subtests except for NO and the Coding Speed (CS) subtest, which is the other speeded test in the ASVAB (84 items in 7 minutes) [1]. The remaining subtests in the ASVAB are power tests with generous time limits. Because the educational level of the 1980 Youth Population is higher than that of military applicants, the group is expected to have higher ASVAB scores.

The reason for the higher NO scores of military examinees lay in the design of the answer sheets, shown in figure 1. The answer spaces for the 1980 Youth Population were circles arranged in five columns of 10 items per column. The circles take more time to fill in than the rectangles that military examinees filled in. In addition, NO items in the test booklets used by both the 1980 Youth Population and military examinees are arranged in the same way as the answer spaces on the military answer sheet—seven columns of seven items each, plus one item in the last column—but the answer sheet used by the 1980 Youth Population was arranged differently. In a special study [3], the differences between the answer sheets were found to account for the differences in NO scores between the 1980 Youth Population and military examinees. The NO and CS scores for the 1980 Youth Population were adjusted to reflect the differences between the testing materials [2]. Thus, only after adjusting the NO scores of the 1980 Youth Population did the 1980 metric provide an accurate basis for describing the aptitudes of military examinees.

The sensitivity of the NO scores to the testing materials resulted in another problem that surfaced before the 1980 score scale was introduced (on 1 October 1984 along with forms 11, 12, and 13 of the ASVAB). In August 1982, new test booklets for form 10 of the ASVAB were introduced, called 10X and 10Y. These were scrambled versions of items introduced in October 1980. The print format in the test booklets was inadvertently changed from the original format for ASVAB 8/9/10, and the result was that NO scores were systematically lowered by about 2 raw score points. Because ASVAB 8/9/10 was scheduled for replacement in October 1984, no changes to the score scale for forms 10X and 10Y were made.

In October 1984, forms 11, 12, and 13 (ASVAB 11/12/13) were introduced along with the 1980 score scale. The accuracy with which ASVAB 11/12/13 was placed on the 1980 metric was evaluated through an Initial Operational Test and Evaluation (IOT&E). Form 8A, which had been administered to the 1980 Youth Population, was administered along with ASVAB 11/12/13 to military applicants in October and November 1984. The results showed that the NO scores for ASVAB 11/12/13 were systematically lower than for ASVAB 8A. This difference implied that the scores for applicants taking ASVAB 11/12/13 were low compared to the 1980 score scale, and their ability was underestimated. In July 1986, an adjustment was made to the AFQT scores that compensated for the lowered NO scores, and the AFQT scores were raised by approximately 2 percentile score points. This problem with the NO scores arose because the type font for printing ASVAB 11/12/13 was inadvertently changed from that used with the test booklets used in the initial calibration of ASVAB 11/12/13.

Answer spaces for the 1980 Youth Population

PART 5 — NUMERICAL OPERATIONS

1 Ⓐ Ⓑ Ⓒ Ⓓ	11 Ⓐ Ⓑ Ⓒ Ⓓ	21 Ⓐ Ⓑ Ⓒ Ⓓ	31 Ⓐ Ⓑ Ⓒ Ⓓ	41 Ⓐ Ⓑ Ⓒ Ⓓ
2 Ⓐ Ⓑ Ⓒ Ⓓ	12 Ⓐ Ⓑ Ⓒ Ⓓ	22 Ⓐ Ⓑ Ⓒ Ⓓ	32 Ⓐ Ⓑ Ⓒ Ⓓ	42 Ⓐ Ⓑ Ⓒ Ⓓ
3 Ⓐ Ⓑ Ⓒ Ⓓ	13 Ⓐ Ⓑ Ⓒ Ⓓ	23 Ⓐ Ⓑ Ⓒ Ⓓ	33 Ⓐ Ⓑ Ⓒ Ⓓ	43 Ⓐ Ⓑ Ⓒ Ⓓ
4 Ⓐ Ⓑ Ⓒ Ⓓ	14 Ⓐ Ⓑ Ⓒ Ⓓ	24 Ⓐ Ⓑ Ⓒ Ⓓ	34 Ⓐ Ⓑ Ⓒ Ⓓ	44 Ⓐ Ⓑ Ⓒ Ⓓ
5 Ⓐ Ⓑ Ⓒ Ⓓ	15 Ⓐ Ⓑ Ⓒ Ⓓ	25 Ⓐ Ⓑ Ⓒ Ⓓ	35 Ⓐ Ⓑ Ⓒ Ⓓ	45 Ⓐ Ⓑ Ⓒ Ⓓ
6 Ⓐ Ⓑ Ⓒ Ⓓ	16 Ⓐ Ⓑ Ⓒ Ⓓ	26 Ⓐ Ⓑ Ⓒ Ⓓ	36 Ⓐ Ⓑ Ⓒ Ⓓ	46 Ⓐ Ⓑ Ⓒ Ⓓ
7 Ⓐ Ⓑ Ⓒ Ⓓ	17 Ⓐ Ⓑ Ⓒ Ⓓ	27 Ⓐ Ⓑ Ⓒ Ⓓ	37 Ⓐ Ⓑ Ⓒ Ⓓ	47 Ⓐ Ⓑ Ⓒ Ⓓ
8 Ⓐ Ⓑ Ⓒ Ⓓ	18 Ⓐ Ⓑ Ⓒ Ⓓ	28 Ⓐ Ⓑ Ⓒ Ⓓ	38 Ⓐ Ⓑ Ⓒ Ⓓ	48 Ⓐ Ⓑ Ⓒ Ⓓ
9 Ⓐ Ⓑ Ⓒ Ⓓ	19 Ⓐ Ⓑ Ⓒ Ⓓ	29 Ⓐ Ⓑ Ⓒ Ⓓ	39 Ⓐ Ⓑ Ⓒ Ⓓ	49 Ⓐ Ⓑ Ⓒ Ⓓ
10 Ⓐ Ⓑ Ⓒ Ⓓ	20 Ⓐ Ⓑ Ⓒ Ⓓ	30 Ⓐ Ⓑ Ⓒ Ⓓ	40 Ⓐ Ⓑ Ⓒ Ⓓ	50 Ⓐ Ⓑ Ⓒ Ⓓ

Answer spaces for military examinees

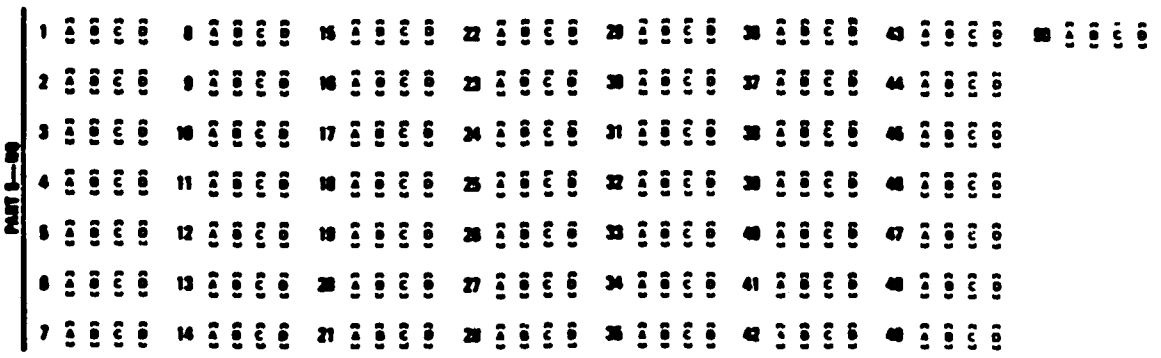


FIG. 1: ANSWER SPACES FOR THE NUMERICAL OPERATIONS SUBTEST USED WITH THE 1980 YOUTH POPULATION AND MILITARY EXAMINEES

The focus of attention from the Joint Services Selection and Classification Working Group has been on problems with NO scores that arise from the design and print format of the testing materials. These problems have, in principle, been resolved. Future versions of the ASVAB test booklets will be printed in a format identical with that used for the 1980 Youth Population. If size, shape, and darkness of the type font and answer spaces were the only problems with the NO subtest, concerns about their accuracy would have been put to rest. Problems with the NO subtest, however, are more pervasive, as discussed in the next section.

INHERENT DEFECTS OF THE NO SUBTEST

The inherent defects of the NO subtest stem from the fact that the items are easy to answer, and the time required to answer them is a major determiner of the test score. Examinees can be easily taught the answers, as most items involve the addition, subtraction and multiplication of all possible pairs of numbers from 1 through 12 and all division problems from 1 through 12 that have whole numbers as answers. Recruiters can easily teach applicants how to do these problems. In addition, examinees can be instructed on how to take the NO subtest. Some rules are: always know which item is next and where to record the answer by keeping your finger on the item being answered; keep your pencil on the answer space corresponding to the item you are working on; make a single stroke to record the answer; work as fast as possible; and keep your attention focused on the problems. These techniques are legitimate strategies for improving NO scores. Teaching examinees these techniques can be a normal part of the preparation, just as taking the practice items in the ASVAB Information Pamphlet helps prepare examinees. An additional strategy is for examinees to continue working for a brief period after time is called. On average, one NO item takes about 5 seconds to complete; completing two extra items is hardly noticeable. These test-taking strategies are helpful when taking any speeded test. Through experience, people have learned these techniques, and perhaps others, that can raise speeded test scores.

Raising the NO scores through test-taking strategies would not have an adverse impact on personnel decisions if two conditions were satisfied. One is that all applicants for enlistment have an equal opportunity to learn the strategies, and the second is that examinees in the 1980 Youth Population have the same opportunity to learn and use the same strategies. The extent to which all applicants use the same test-taking strategies is not known. What is known is that the 1980 Youth Population generally did not know or follow

these strategies. When the 1980 Youth Population was tested, few individuals were aware of how sensitive speeded tests are to these influences. The examinees had no preparation on how to take NO. In fact, during the first months that ASVAB 8/9/10 was in use, the military examinees probably had little preparation on how to improve their scores.

The first analysis of test-taking strategies shows the effects of practice on NO scores. The Army Research Institute (ARI) has reported two studies showing that retesting on the ASVAB increases NO scores more than scores on the power tests. In one study [4], 53 examinees from a Job Corps Training Center were retested with five different forms of the ASVAB on successive days. The mean scores for NO and Math Knowledge (MK), a power test that has a generous time limit, for each of the five days are shown in figure 2. The NO scores show a fairly steady increase, from 38 in the first session to 44 in the fifth; the increase is over one-half of a standard deviation. By contrast, the MK scores remained relatively stable. The initial mean was 41, and the final mean was 42.

In another study [5], the scores of Army applicants in fiscal year 1981 who failed to qualify for enlistment on their initial ASVAB testing were compared to their scores on a retest. Over 17,000 examinees were included in the analysis. The mean increase in NO, MK, and AR scores on the retest are shown in figure 3. The NO scores increased by 3 subtest standard score points (about one-third of a standard deviation), while the means for the other two math tests, both of which are power tests, increased by less than 1 point. These results show that NO scores can be raised through retesting.

Retesting, however, does not automatically increase NO scores. Figure 4 shows the effects of voluntary retesting at military testing stations, when scores increase, versus mandatory testing, when scores go down. Voluntary retesting occurs when applicants want to improve their test scores, either because they failed to qualify for enlistment or for a desirable option, such as a bonus. Mandatory retesting occurs when an examinee's AFQT score is excessively high compared to other ASVAB subtests not in the AFQT. Applicants are more likely to cheat on the AFQT than on other subtests, and an excessively high AFQT score relative to their performance on other selected subtests may indicate cheating. These examinees were tested with another form of the AFQT to verify their aptitude level. For the examinees retested under the mandatory program, the AR scores decreased, perhaps because they tended to be coached on the specific items in the first AFQT they took. The MK scores remained essentially unchanged.

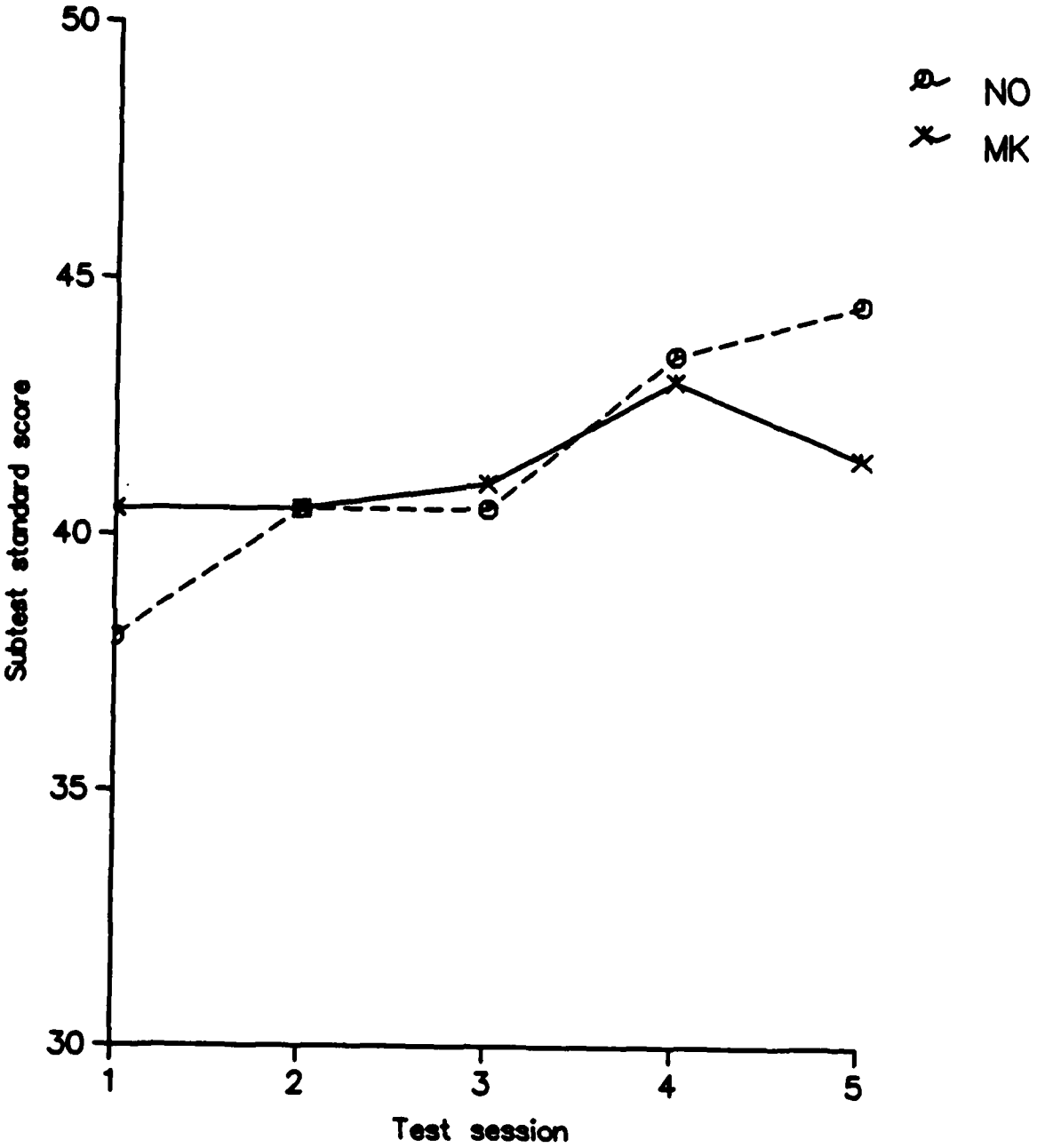


FIG. 2: MEAN NO AND MK SCORES FOR EXAMINEES TESTED FIVE TIMES

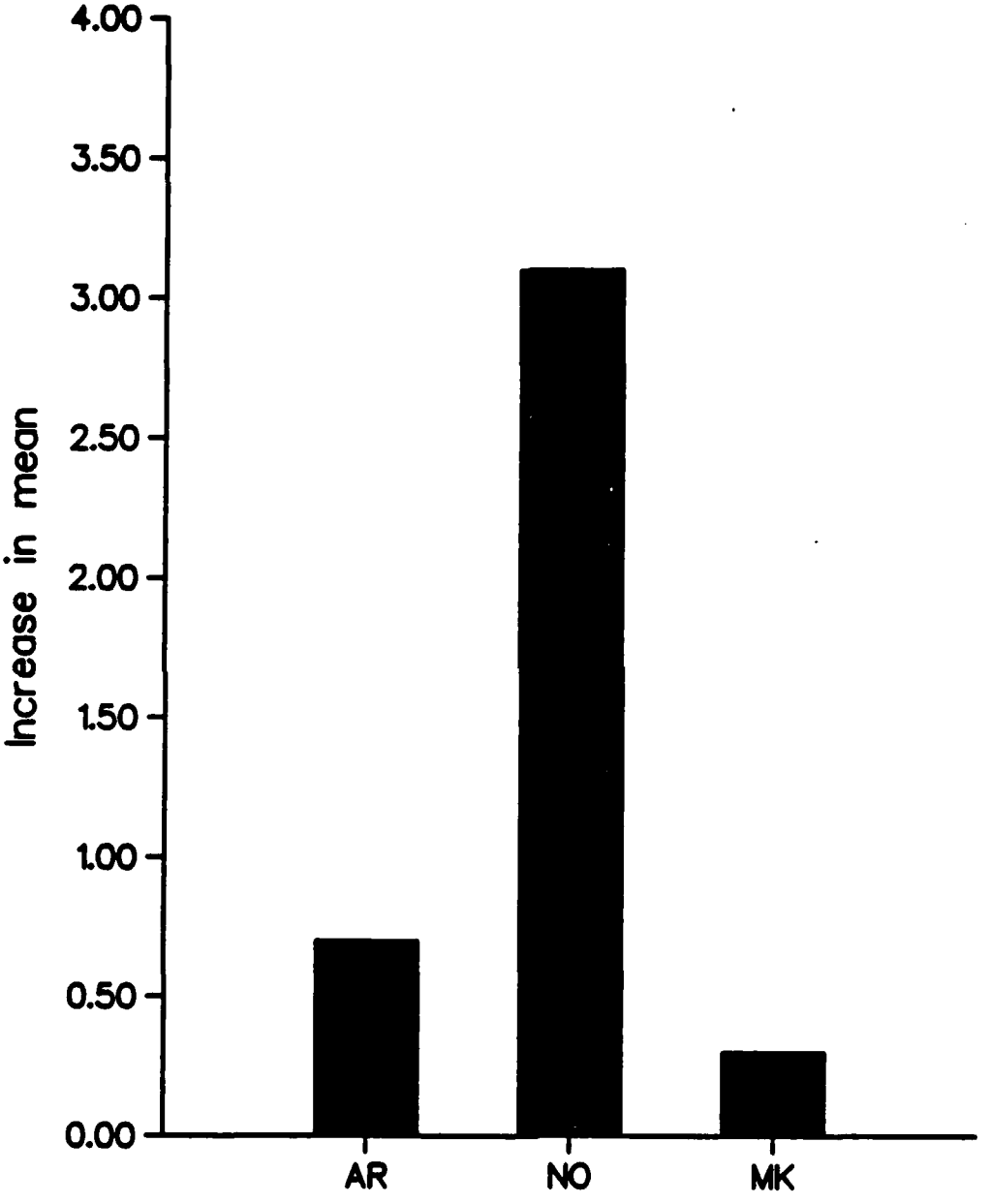


FIG. 3: EFFECTS OF VOLUNTARY RETESTING ON ARMY APPLICANTS

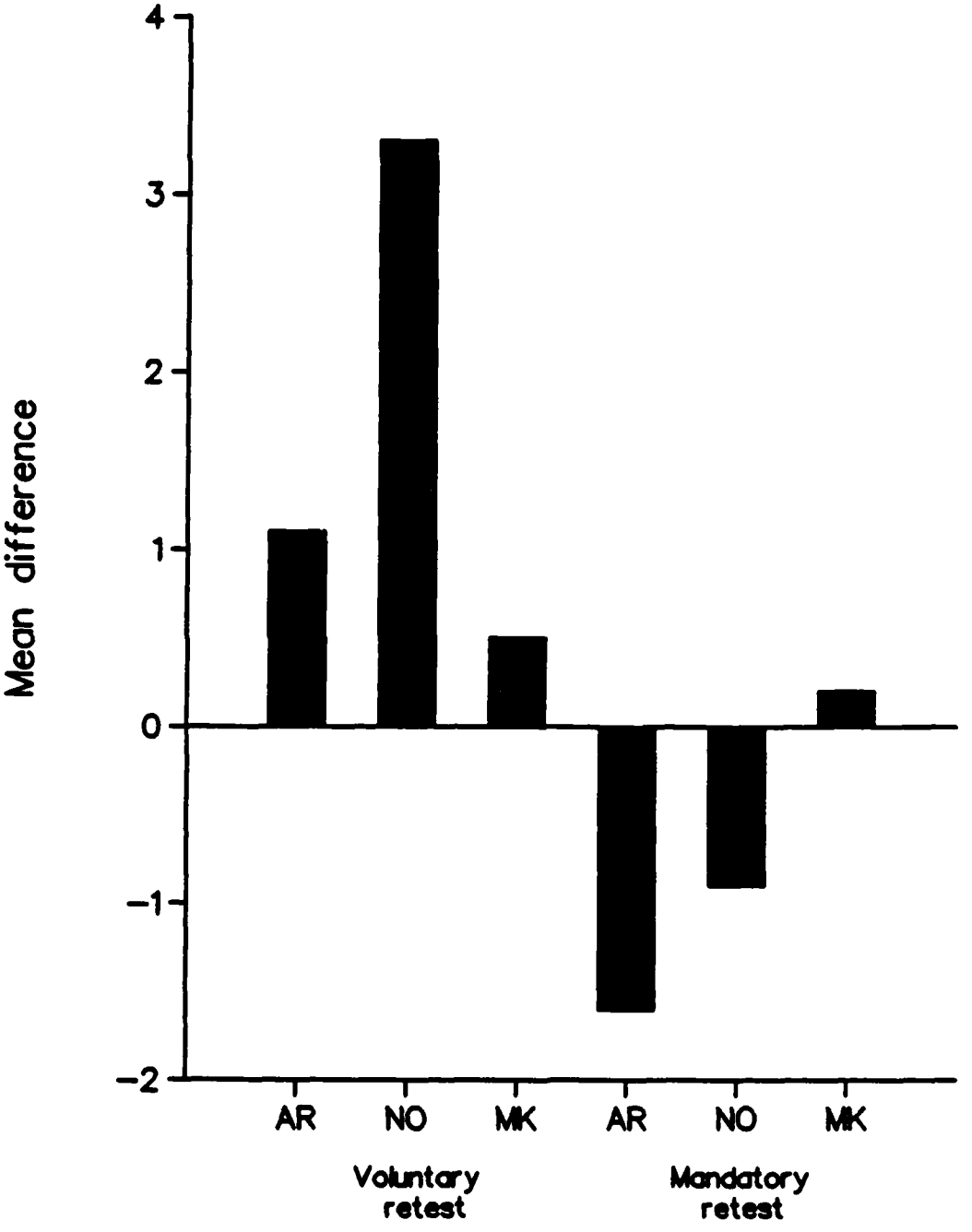


FIG. 4: EFFECT OF RETESTING ON SUBTEST SCORES OF DOD APPLICANTS

The second analysis of test-taking strategies compares the NO scores for examinees in three groups: the 1980 Youth Population; the 1980 IOT&E, when NO was first made part of the AFQT and before the strategies were widely practiced; and the 1984 IOT&E, when the strategies were fully known and practiced. For each group, the mean NO score conditional on the AR scores was computed; that is, the mean NO score was computed for all people in each group that had the same AR score, for each AR score in turn. The 1984 IOT&E sample was restricted to examinees who enlisted for active duty; people applying for enlistment in the reserve components were deleted. Active-duty and reserve-components applicants could not be distinguished in the 1980 IOT&E sample. By computing mean NO and MK scores at each AR score, differences in ability among the three groups were statistically controlled, and the differences in NO scores more nearly represent the effect of test-taking strategies. For comparison, the MK means at each AR score were also computed. Because MK was not part of the AFQT and because it is not as susceptible to test-taking strategies, the MK scores for people with the same AR score are expected to remain stable for the three groups. Separate analyses were completed for males and females because females score better than males on speeded tests, and comparison of the three groups would be distorted by differing proportions of females. Analyses were also performed for racial groups.

The mean NO and MK scores at each AR score are shown in figure 5 for males, figure 6 for females, figure 7 for white males, figure 8 for black males, figure 9 for white females, and figure 10 for black females. In every case the results are similar; compared to the 1980 Youth Population, the NO scores for the 1984 IOT&E group show a large amount of inflation, and those for the 1980 IOT&E group show a small amount of inflation. Over the years NO scores have increased, especially for people with low AR scores. These people on average have NO scores from 5 to 7 points higher than people in the 1980 Youth Population. The differences are most pronounced for people with AR scores below 15, which is in the below-average range. The differences tend to be more pronounced for blacks than for whites.

These results demonstrate that NO scores have increased since NO was made a part of the AFQT in October 1980. Although the NO scores in the 1980 IOT&E were somewhat higher than the 1980 Youth Population for people with AR scores below 15, the big jump took place later, as evidenced by the high NO scores for the 1984 IOT&E group. In contrast, the MK scores remained stable.

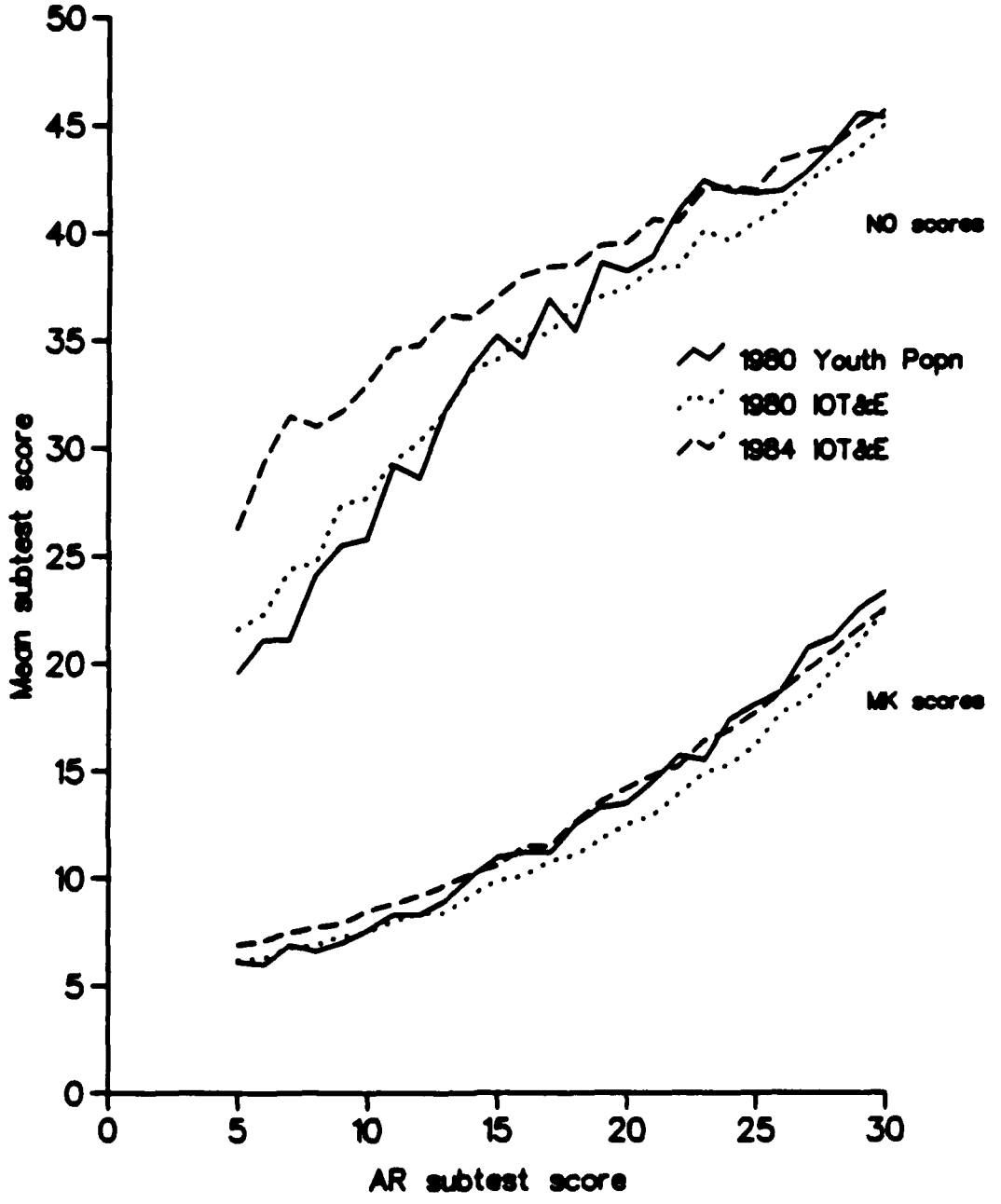


FIG. 5: MEAN NO AND MK SCORES SHOWN BY AR SCORE FOR MALES

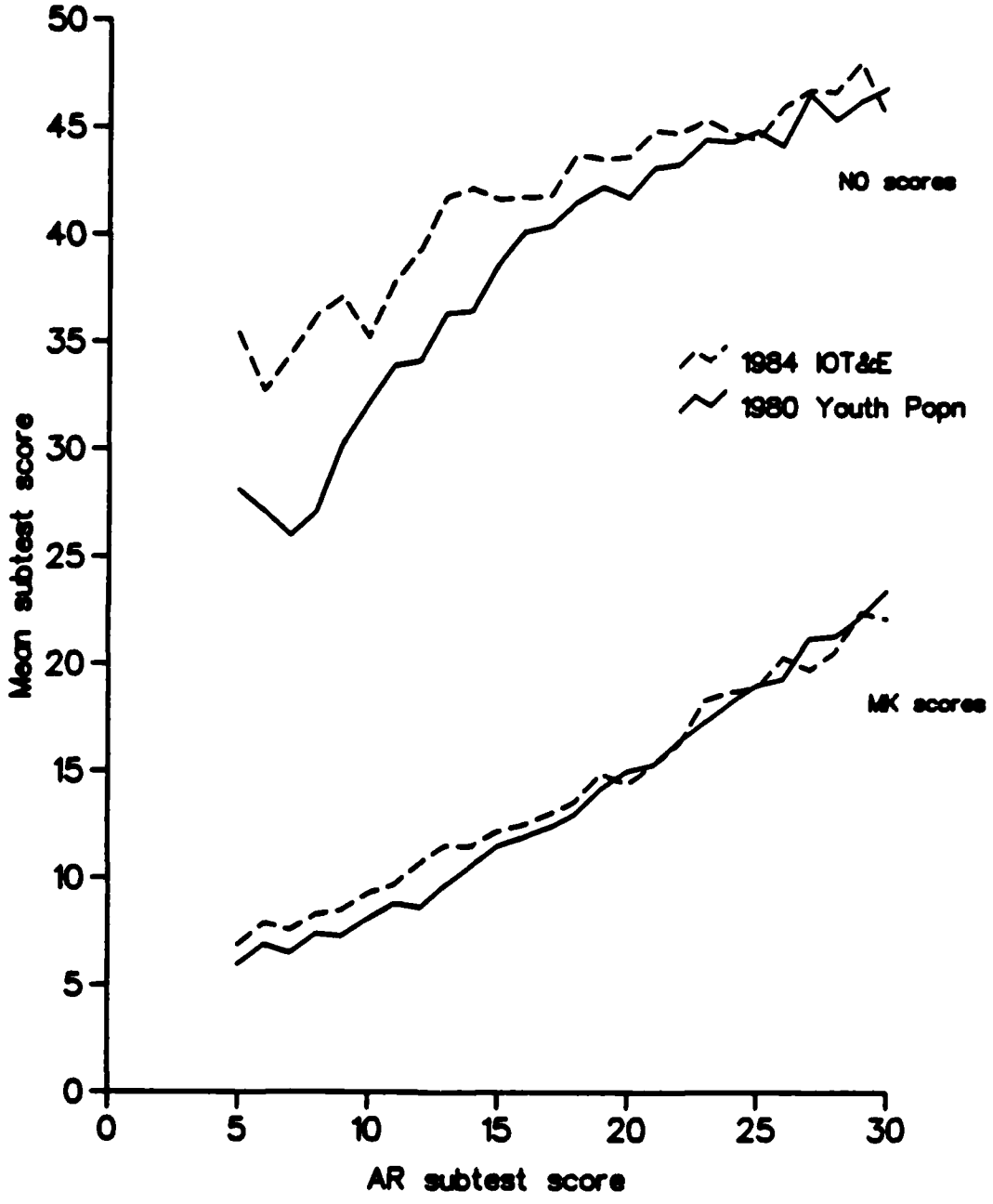


FIG. 6: MEAN NO AND MK SCORES SHOWN BY AR SCORE FOR FEMALES

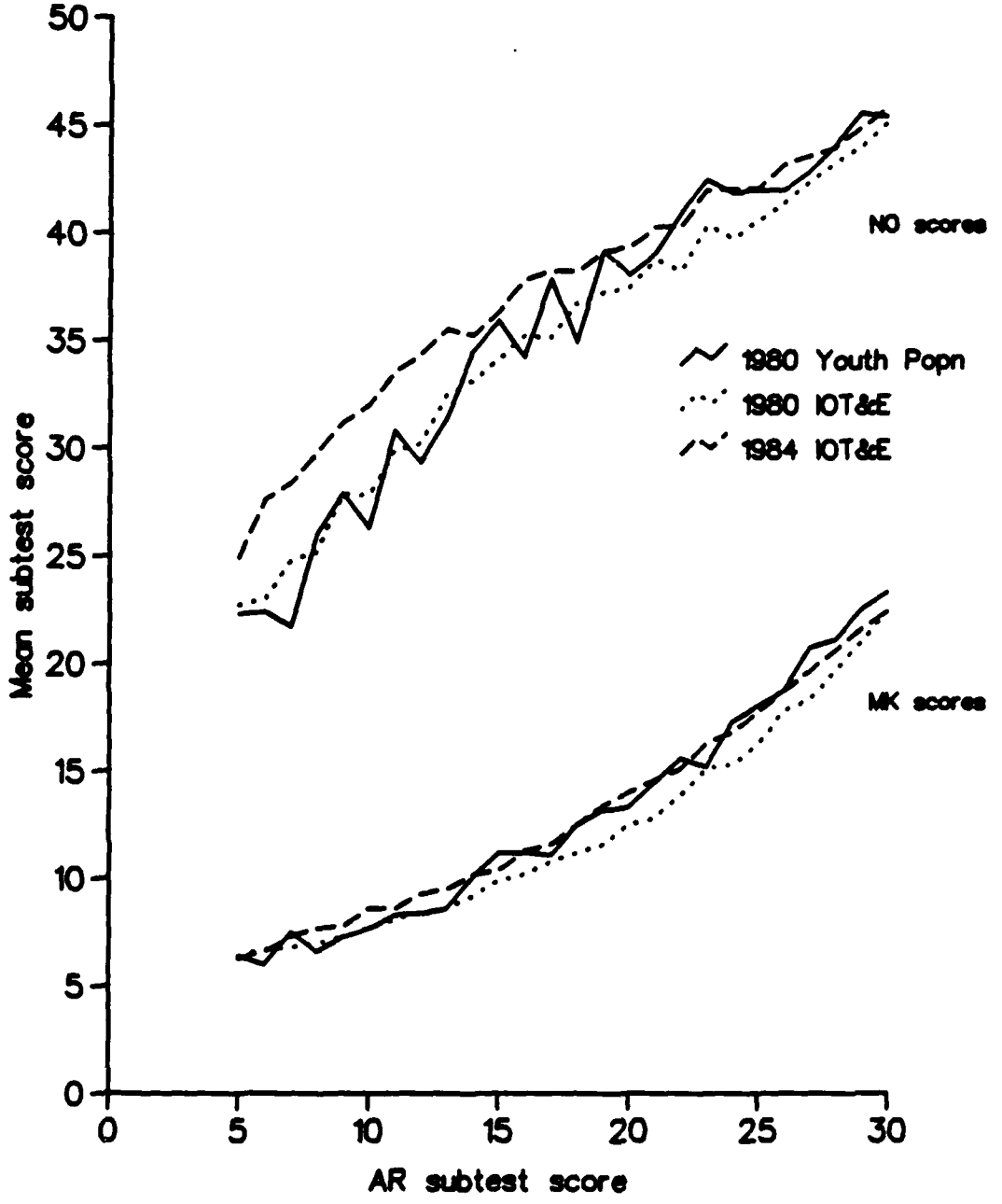


FIG. 7: MEAN NO AND MK SCORES SHOWN BY AR SCORE FOR WHITE MALES

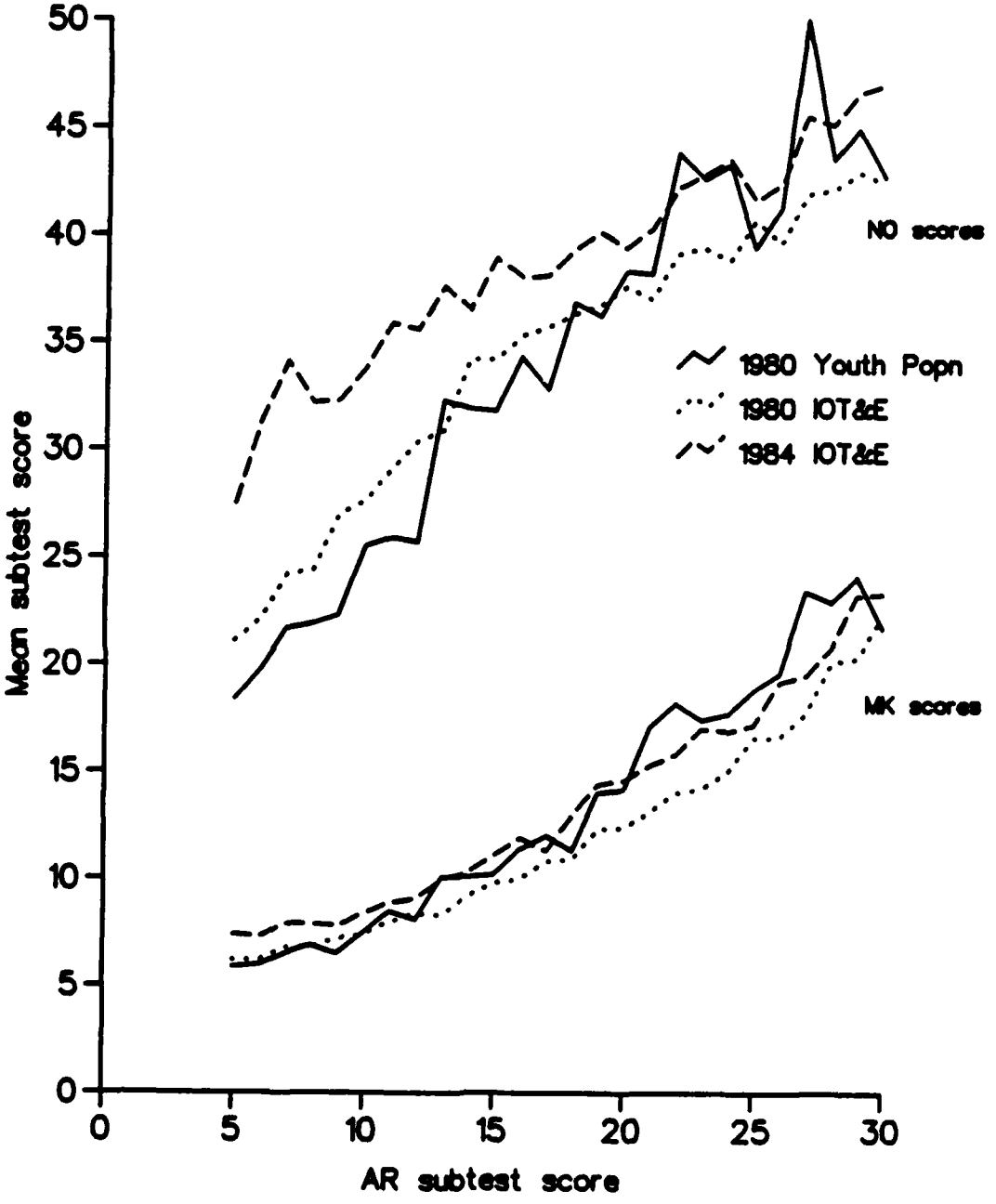


FIG. 8: MEAN NO AND MK SCORES SHOWN BY AR SCORE FOR BLACK MALES

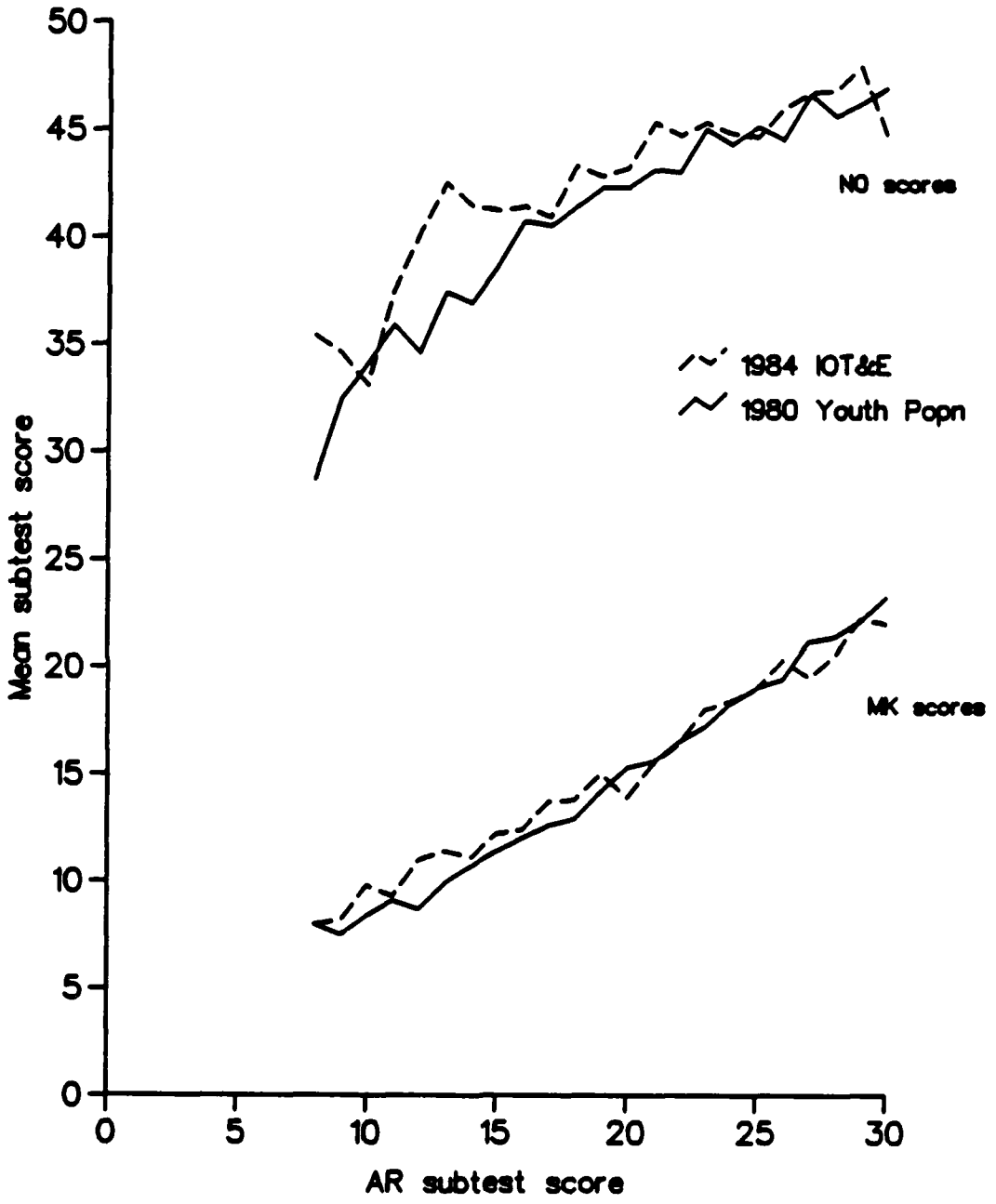


FIG. 9: MEAN NO AND MK SCORES SHOWN BY AR SCORE FOR WHITE FEMALES

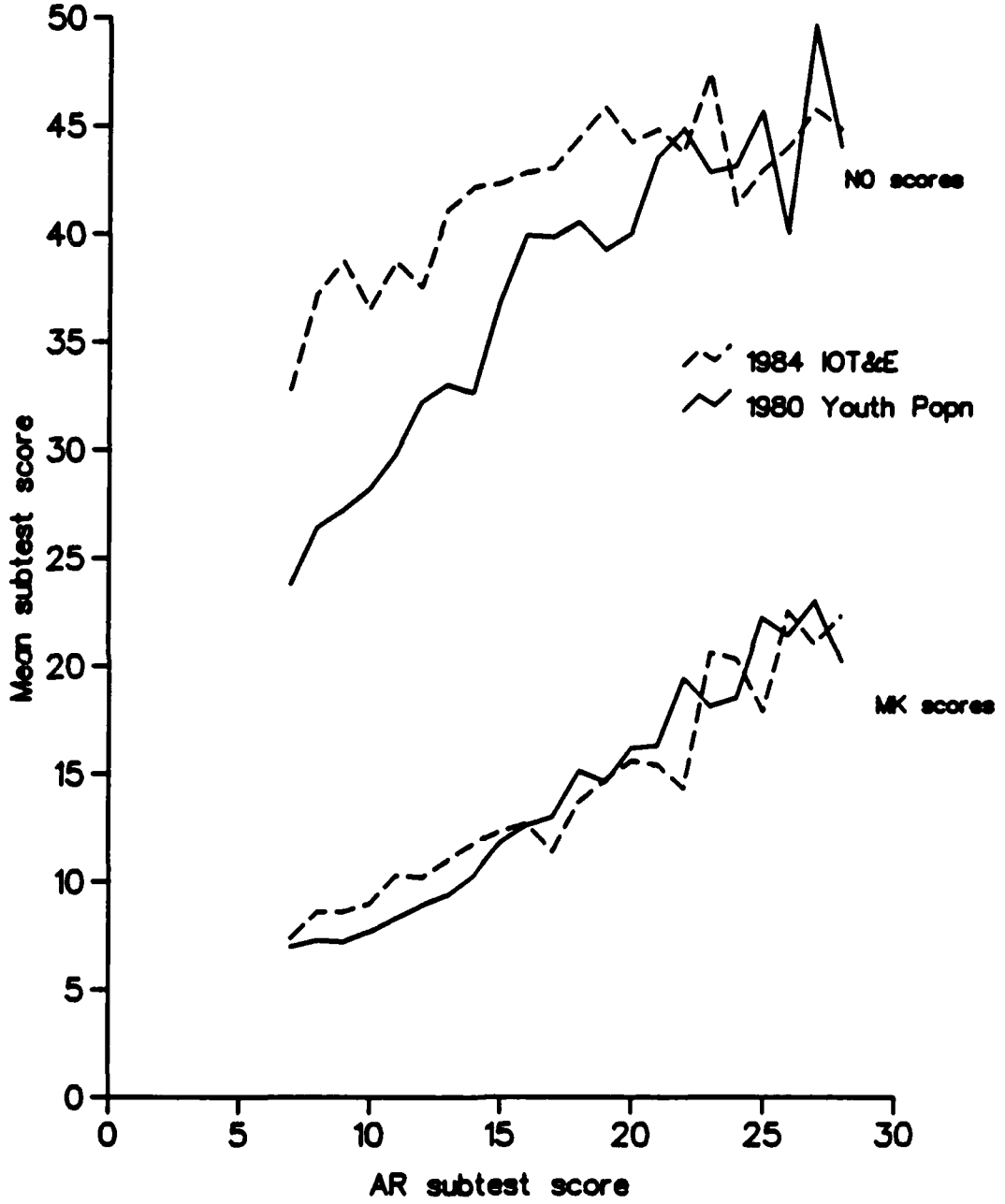


FIG. 10: MEAN NO AND MK SCORES SHOWN BY AR SCORE FOR BLACK FEMALES

Additional effects of test-taking strategies on the NO scores were evaluated in the 1984 IOT&E group. Some examinees were enlisting for active duty and some for the reserve components (Guard and Reserves). Separate NO means conditional on AR scores were computed in the two groups. The results are shown in figure 11 for males and figure 12 for females. The amount of inflation is noticeably higher for the active-duty applicants. The suggestion is that the active-duty applicants were more aware of and followed the test-taking strategies to a greater degree than did the reserve-components applicants. As noted earlier, applicants for the reserve components were not included in the previous results for the 1984 IOT&E in figures 5 through 10.

Another analysis was based on examinees in the 1984 IOT&E who marked on their answer sheet that they were being retested on the ASVAB. The answer sheet has a section labeled "TEST TYPE," with the response options: INITIAL, RET-MEPS, RET-REC, and VERIFIC. Of the 96,137 examinees in the 1984 IOT&E who were applying for active duty, 79,326 marked the INITIAL space; none marked RET-MEPS; 3,996 marked RET-REC; and 12,815 marked VERIFIC. The instructions for completing the test-type section in the ASVAB Administration Manual are vague, and each examinee or test administrator can uniquely interpret the labels. Presumably, INITIAL means first test, and the other responses mean that the examinee is retesting. These numbers are discrepant from the normal amount of retesting 4,000-5,000 per month who voluntarily retest, and 150-200 per month who are retested to verify (VERIFIC) their AFQT score. Separate analyses were performed on the examinees who marked INITIAL and those who marked the other spaces. The results are shown in figure 13 for males and figure 14 for females.

People who marked INITIAL, and presumably were taking the ASVAB for the first time, had markedly less inflation of their NO scores than those who marked the other spaces, and presumably were being retested. (There were no differences in the conditional NO means between the examinees who marked RET-REC and VERIFIC.) These results are consistent with those shown earlier (figure 3) that people who voluntarily retest improve their NO scores more than their scores on power tests, which have generous time limits.

The amount of inflation of NO scores in the 1984 IOT&E group is based on all applicants regardless of their response in the test-type section; that is, test type (initial or retest) was ignored. Applicants for the reserve components were not included, except in figures 11 and 12. The amount of inflation, thus, is based on a group of applicants that represents a normal flow through the

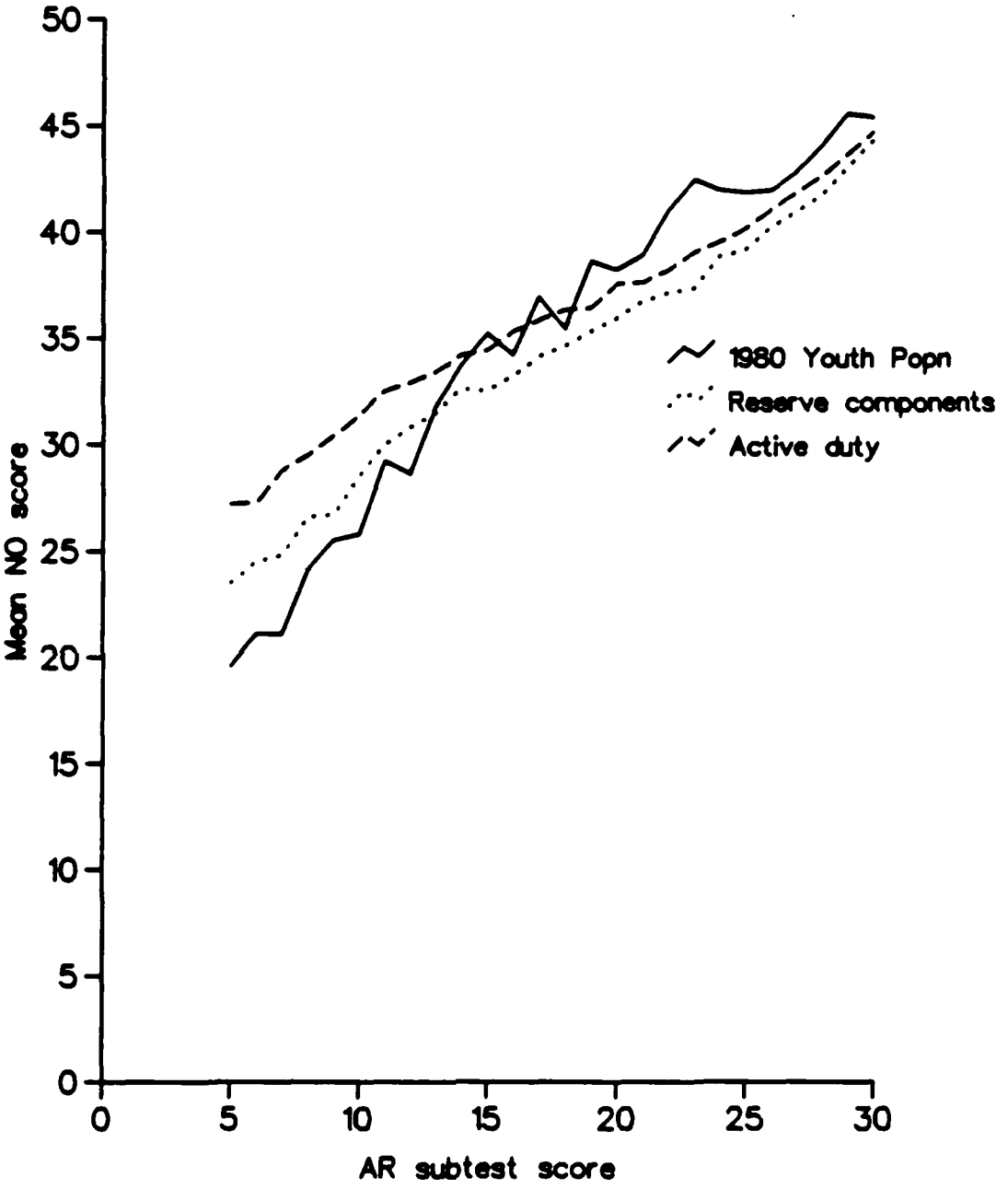


FIG. 11: MEAN NO SCORE SHOWN BY AR SCORE FOR ACTIVE DUTY VS. RESERVE (MALES)

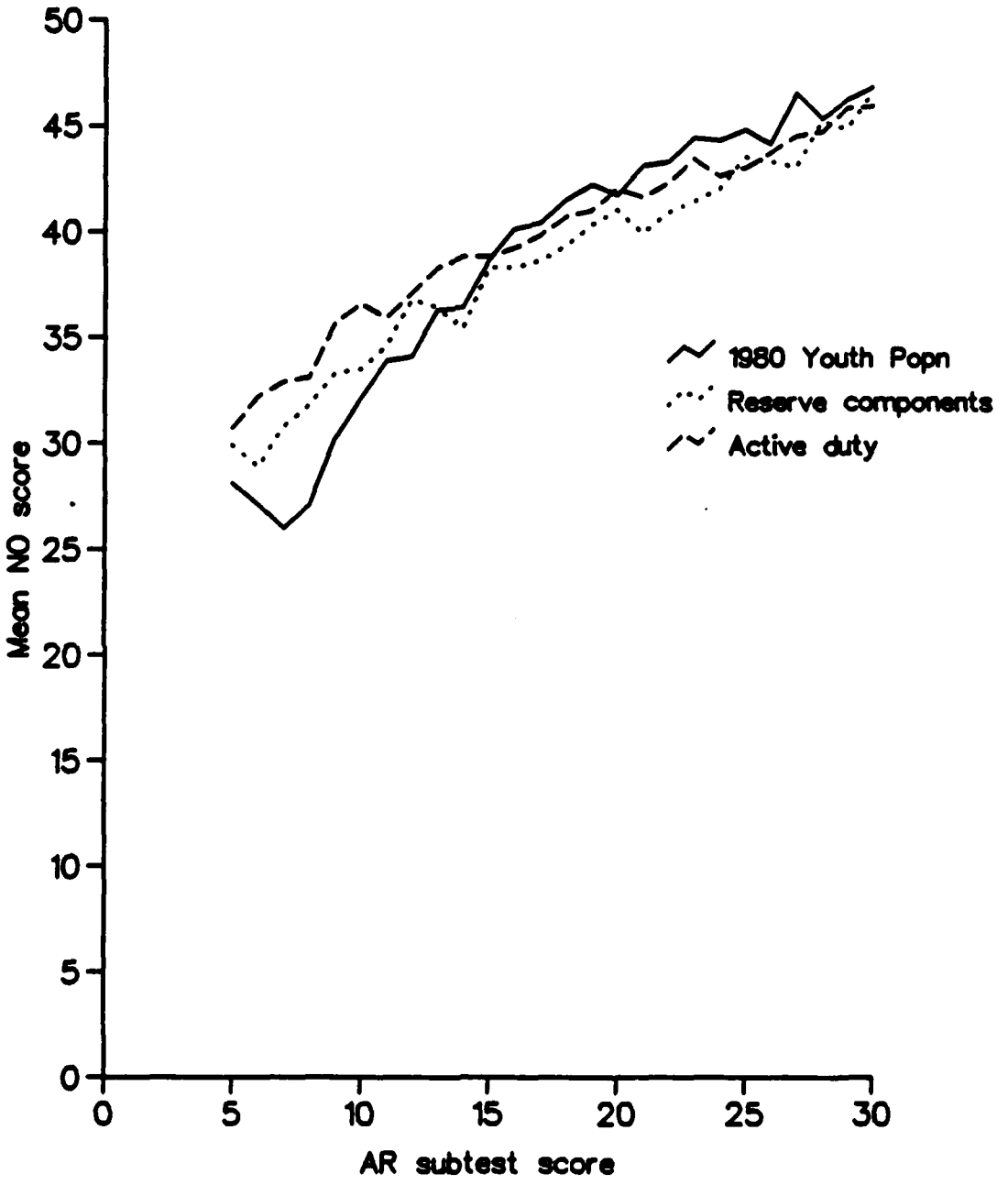


FIG. 12: MEAN NO SCORE SHOWN BY AR SCORE FOR ACTIVE DUTY VS. RESERVE (FEMALES)

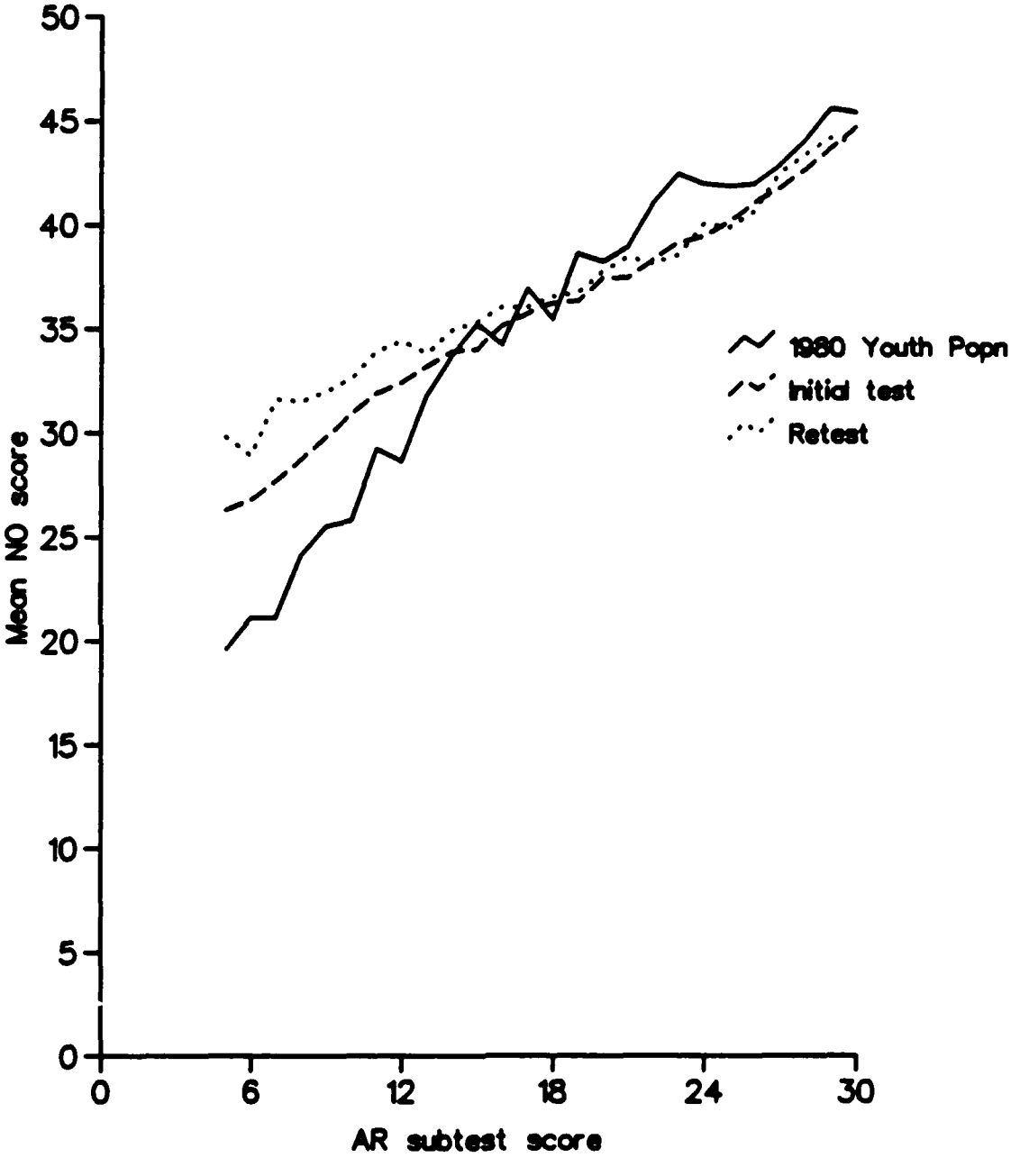


FIG. 13: EFFECTS OF RETESTING ON NO SCORES (MALES)

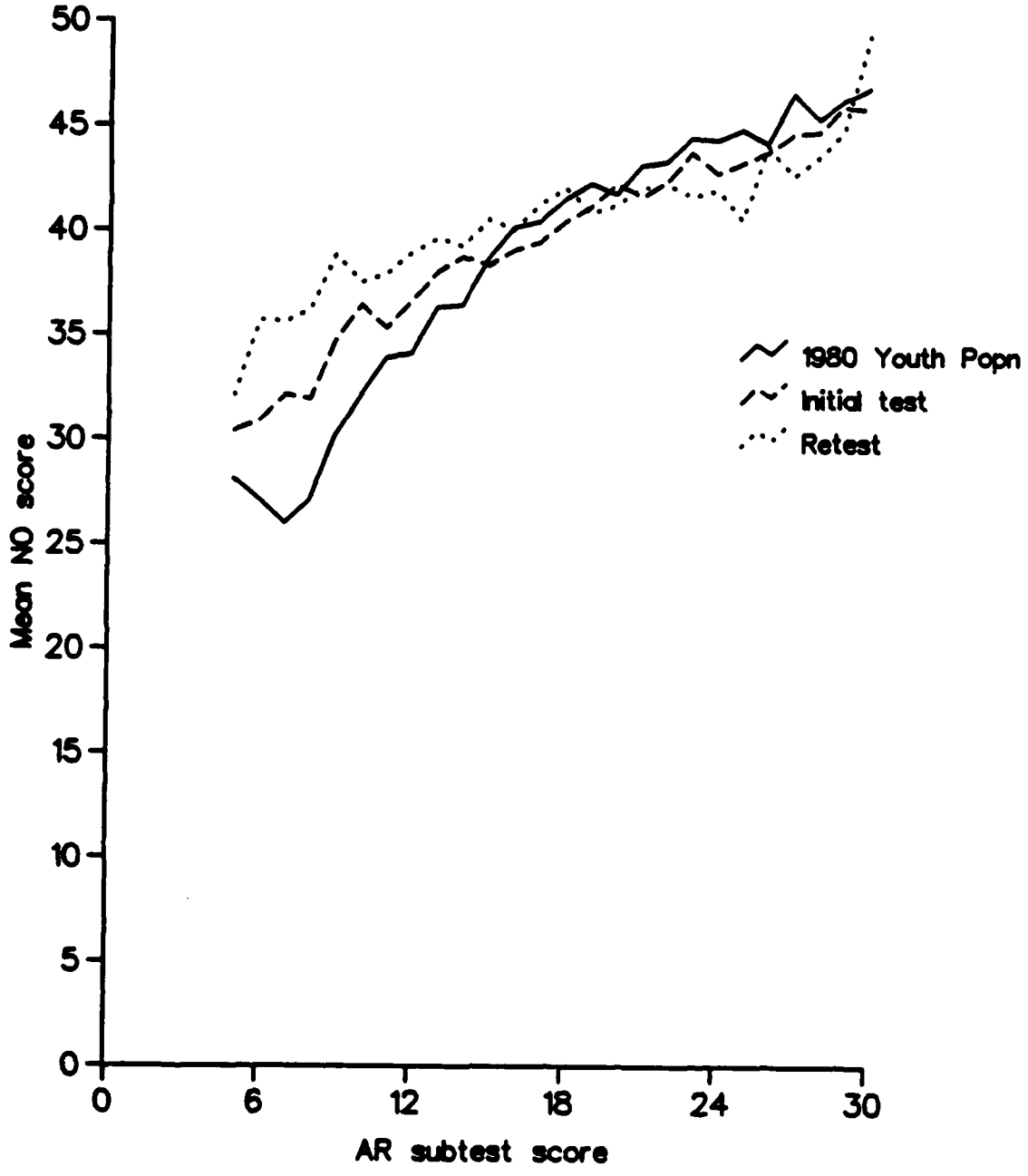


FIG. 14: EFFECTS OF RETESTING ON NO SCORES (FEMALES)

testing stations. The amount of inflation reported here is the expected amount in normal groups of applicants. Because a purpose of this analysis is to evaluate the amount of inflation in operational test scores, the people who are retested with the ASVAB should be included.

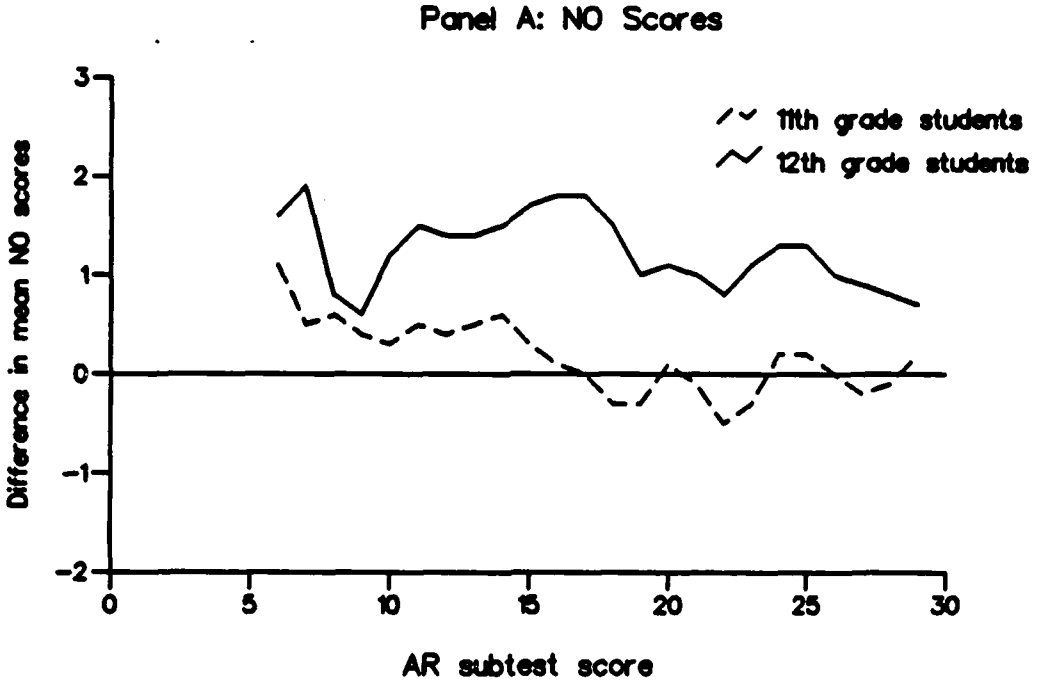
Form 14 of the ASVAB is administered to high school students nationwide as part of the Student Testing Program. The test scores can be used to help determine qualification for enlistment, and the names of examinees provide valuable leads to recruiters. The students can mark on the answer sheets their postsecondary-school interests. Options include military service and vocational-technical training. The change in NO scores between grades 11 and 12 was compared for these two groups of examinees who were tested during academic year 1984-85. The change in MK scores was also evaluated. The conditional NO and MK means, controlling on AR were computed for each group (postsecondary interests and grade). Because the number of cases for some means is relatively small (a minimum of 100 cases), the means were smoothed using the 3-point moving average, weighted .25, .50, and .25.

The differences in conditional means between grades 11 and 12 for the students interested in military service or in vocational-technical (voc/tech) training are shown in figure 15. Only one difference is large—the difference in NO scores between 11th and 12th grade students interested in the military services is larger than the other sets of differences (NO scores for students interested in vocational-technical training and the MK scores). Apparently the 12th grade students interested in the military learned and used the test-taking strategies to a greater extent than the other groups of students. The conditional MK means, as expected, did not vary appreciably among the groups.

The next question is how much have the NO scores been inflated compared to the 1980 Youth Population, which was used to construct the 1980 metric, and what is the impact of these inflated scores on qualification rates? These issues are addressed in the next section.

IMPACT OF INFLATED NO SCORES ON QUALIFICATION RATES

Evaluating the impact of the inflated NO scores on qualification rates involves two steps. The first is to quantify the amount of inflation of AFQT scores. The second is to compute the qualification rates of people in the 1984



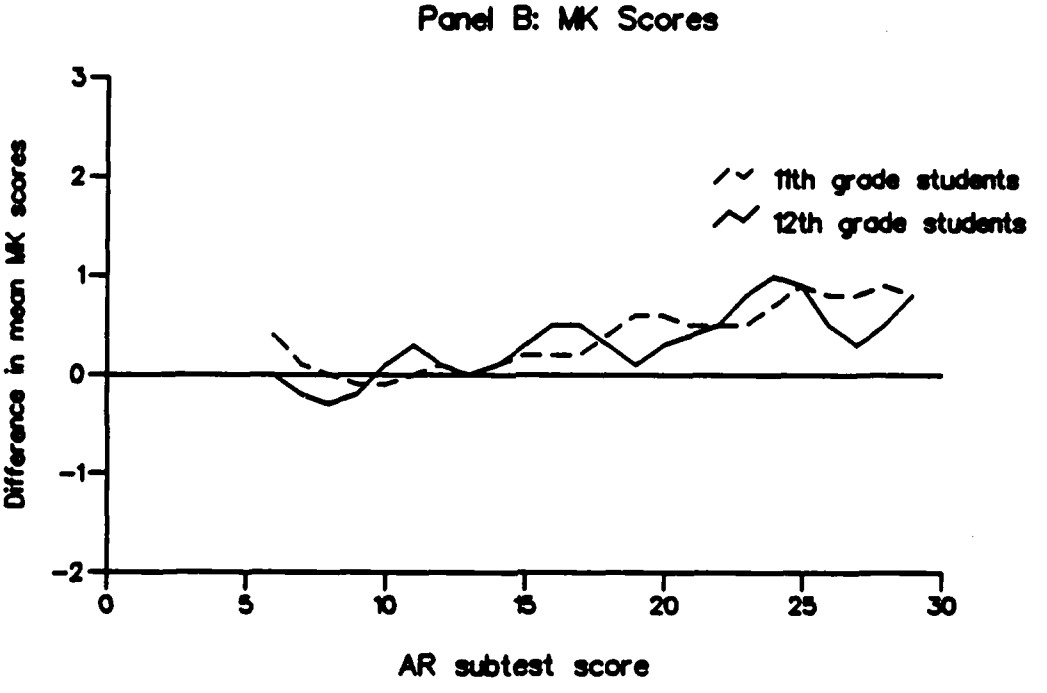


FIG. 15: DIFFERENCE IN NO AND MK SCORES FOR STUDENTS INTERESTED IN VOC/TECH TRAINING AND MILITARY SERVICE

IOT&E on three versions of the AFQT. One is the operational AFQT scores, which includes the inflated NO scores; the second is the adjusted AFQT scores, where the NO scores have been lowered to compensate for the inflation; and the third is the proposed AFQT scores, where NO has been replaced by the MK subtest. The Joint Services Selection and Classification Working Group had recommended that MK replace NO in the AFQT. Because the decision was not unanimous, this analysis was undertaken to evaluate how qualification rates would be affected by retaining or deleting NO in the AFQT.

The amount of inflation in the NO scores was determined through an analytic equipercentile equating procedure. In this technique scores that have the same cumulative frequency are said to be equivalent. The equating determined the NO score in the 1984 IOT&E group that was equivalent to a given NO score in the 1980 Youth Population. Separate equatings were performed for males and females. Details of the equating are presented in appendix A. The equivalent scores are shown in table 2, and the amount of inflation is portrayed in figure 16 for males and figure 17 for females. The maximum inflation is 8 points for males with NO scores from 9 through 14 in the 1984 IOT&E and 7 points for females at NO scores from 8 through 23 in the 1984 IOT&E. The inflation is gradually reduced above these points, until it reaches zero at an NO score of 47 for males and 48 for females. The maximum NO score is 50, which means that the scores for the 1984 IOT&E group are inflated throughout almost the entire range.

The equivalent scores in table 2 were used to adjust the NO scores for people in the 1984 IOT&E group. For example, an NO score of 20 was reduced to 14 for males and 13 for females, and a score of 36, which is close to the population mean, was reduced to 33 for males and 32 for females. These adjusted, or deflated, NO scores were used to compute a new set of AFQT scores, called adjusted AFQT, used in determining qualification rates for the 1984 IOT&E groups.

Qualification rates were computed for the three versions of the AFQT for people in the 1984 IOT&E population tested with forms 11A, 11B, 12B, 13A, and 13B of the ASVAB.[1] The operational AFQT is defined as VE plus AR plus one-half of the inflated NO scores; these AFQT scores were used in making personnel decisions about the qualification of applicants for enlistment and

1. Form 12A was not included because it is more difficult than the others, and a different adjustment to the NO scores may be required.

TABLE 2

EQUIVALENT NO SCORES FOR THE 1980 YOUTH POPULATION AND THE 1984 IOT&E EXAMINEES

NO score in the 1980 Youth Population	1984 IOT&E examinees	
	Males	Females
1	1-9	1-8
2	10	9
3	11	10
4	12	11
5	13	12
6	14	13
7	14	14
8	15	15
9	16	16
10	17	17
11	18	18
12	19	19
13	19	20
14	20	21
15	21	22
16	22	23
17	23	23
18	24	24
19	24	24
20	25	25
21	25	26
22	26	27
23	27	27
24	28	28
25	29	29
26	30	30
27	31	31
28	32	32
29	33	33
30	34	34
31	35	35
32	35	36
33	36	37
34	37	37
35	38	38
36	39	39
37	40	40
38	41	41
38	41	42
39	42	43
40	42	44
41	43	45
42	44	46
43	45	46
44	45	47
45	46	47
46	47	47
47	47	48
48	48	48
49	49	49
50	50	50

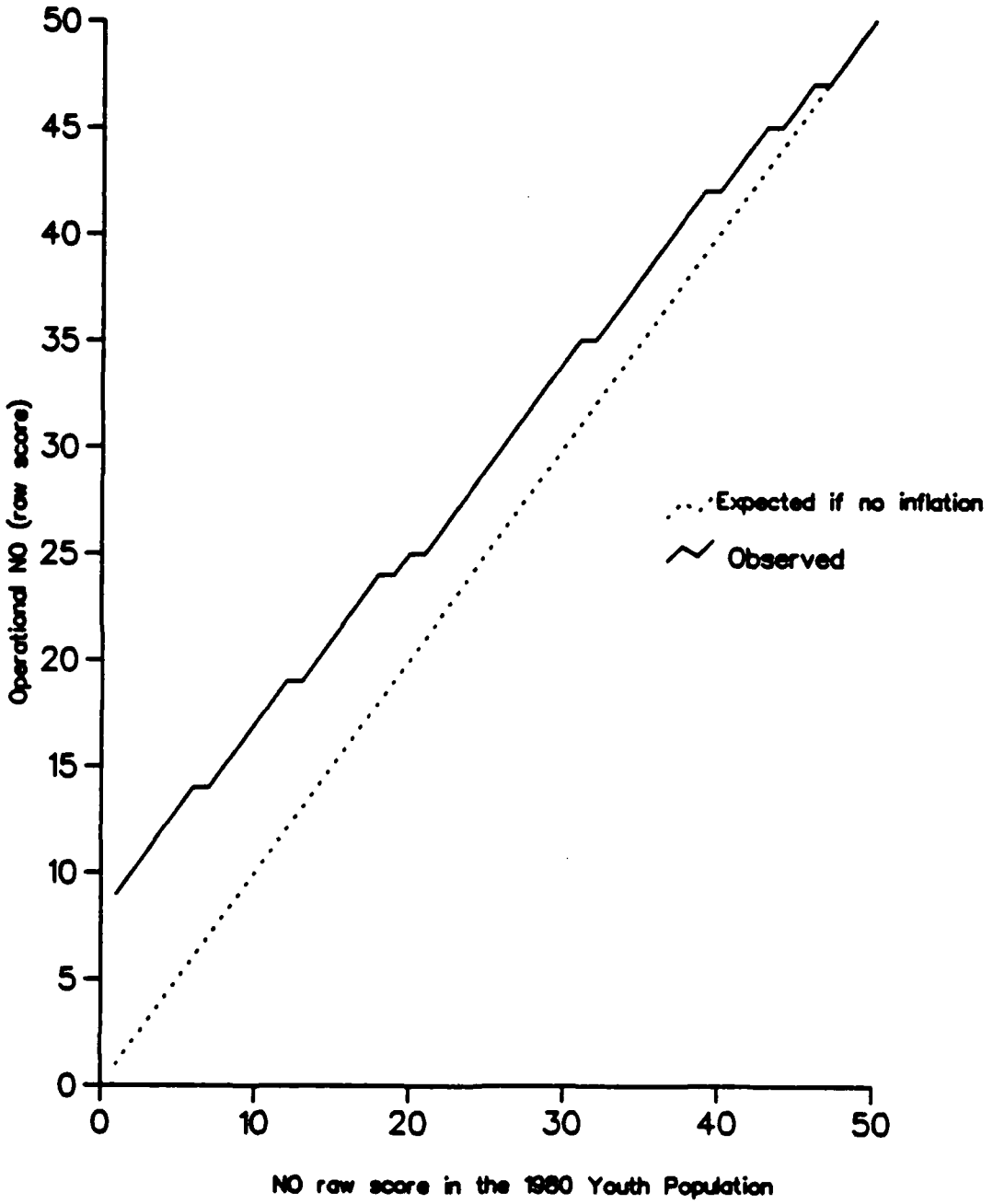


FIG. 16: INFLATION IN NO SCORES FOR MALES

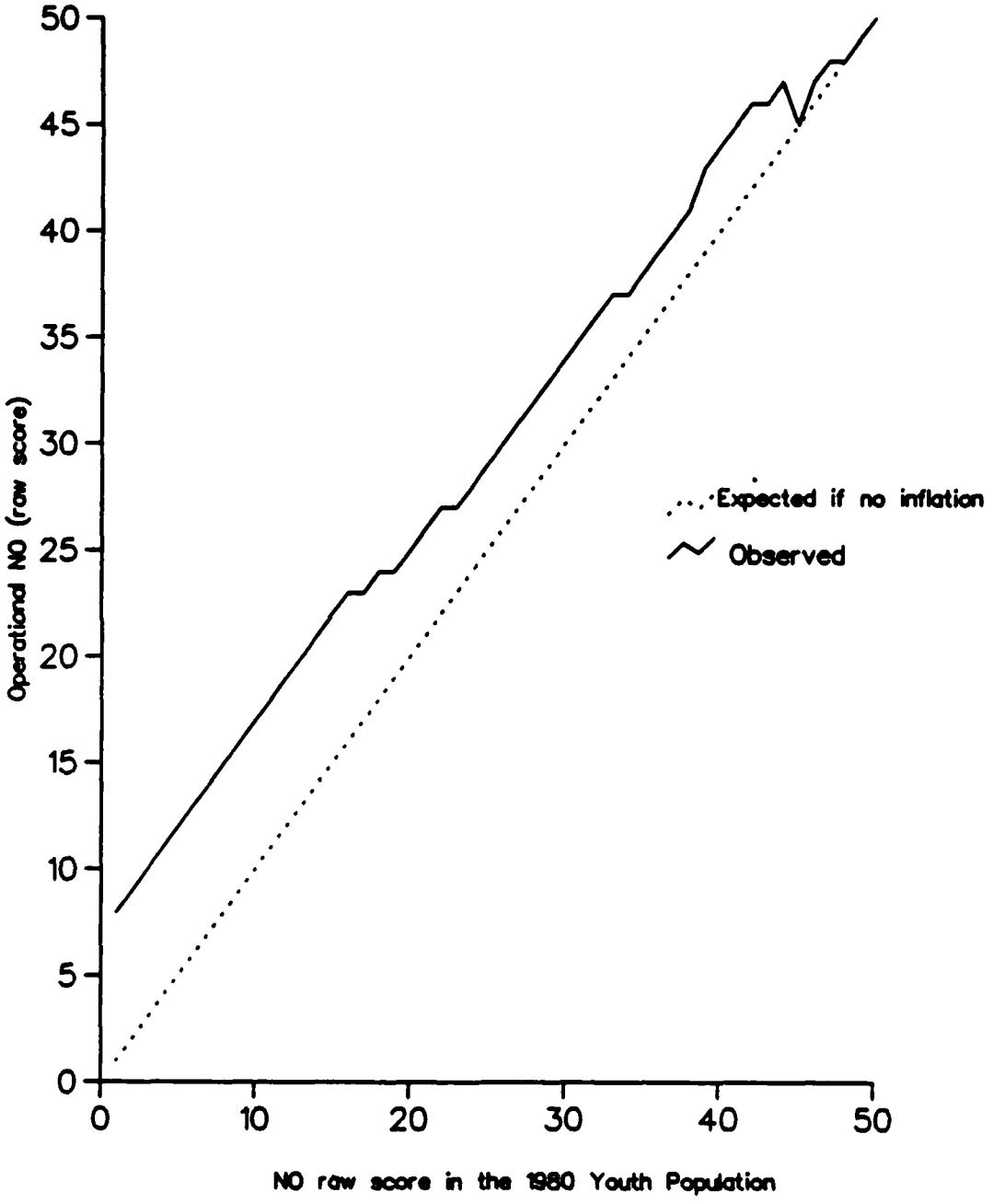


FIG. 17: INFLATION IN NO SCORES FOR FEMALES

special programs, such as bonuses and guaranteed assignments. The qualification rates based on the adjusted AFQT reflect the personnel decisions that would have been made if the NO and AFQT were accurately scaled to the 1980 metric. The proposed AFQT is defined as VE plus AR plus MK. The proposed AFQT scores were placed on the 1980 metric through the equipercentile equating technique. Details are presented in appendix B. Qualification rates on each version of the AFQT were defined as scoring at or above percentile scores of 21, which is sometimes used as the minimum enlistment standard on the AFQT; 31, which is the lower bound for category III, the average range of AFQT scores, and is sometimes used as the minimum enlistment standard on the AFQT; and 50, which is often the minimum qualifying score for bonuses and guaranteed assignments. Thus, three qualification scores on each of three versions of the AFQT were computed for each person in the 1984 IOT&E group who took forms 11A, 11B, 12B, 13A, or 13B.

Only applicants for active duty were included; applicants for the reserve components were excluded. Applicants who retested on the ASVAB were included when computing the qualification rates. Examinees who marked INIT on their answer sheet were analyzed as a separate group. The amount of inflation and qualification rates are shown in appendix C. The results are essentially the same as for the entire group, which includes those who marked that they were being retested.

The qualification rates on each AFQT were computed for the following categories of people in the 1984 IOT&E group:

- Service – Army, Navy, Air Force, Marine Corps and total Department of Defense (DOD)
- Gender – All groupings by race are shown separately for each sex
- Race – blacks, whites, and others.

The qualification rates are shown by service: total DOD in table 3, Army in table 4, Navy in table 5, Air Force in table 6, and Marine Corps in table 7. The qualification rates at percentile scores 31 and 50 for the total DOD examinees are also portrayed in figures 18 through 23 for males, females, white males, black males, white females, and black females, respectively.

TABLE 3

QUALIFYING RATES ON ALTERNATIVE AFQTS
DEPARTMENT OF DEFENSE APPLICANTS

Group	Number of cases[a]	Percent qualifying on AFQT at percentile score: 21 Oper[b]	21 Adj[c]	21 Prop[d]	31 Oper[b]	31 Adj[c]	31 Prop[d]	50 Oper[b]	50 Adj[c]	50 Prop[d]
Males										
Total	53017	88.7	86.6	89.1	74.4	71.2	75.6	46.9	44.0	47.2
White	36616	94.2	92.8	94.8	84.0	81.2	85.2	57.9	54.5	57.9
Black	11323	75.2	71.0	75.2	49.8	45.7	50.6	19.0	17.0	19.5
Other	5078	79.4	76.6	79.9	59.9	57.0	62.5	30.4	28.0	31.6
Females										
Total	9755	92.3	90.4	91.1	78.9	75.6	76.6	49.0	45.8	44.6
White	5978	96.6	95.5	96.2	88.6	86.0	87.2	63.1	59.7	58.1
Black	3001	85.6	82.2	82.2	62.0	57.3	58.0	24.3	21.4	20.7
Other	776	85.2	83.0	86.1	69.7	66.0	66.8	36.1	32.7	32.7

---

a. Applicants for active duty tested October and November 1984.
b. Operational AFQT score, includes inflated NO scores.
c. Adjusted AFQT score, includes deflated NO scores.
d. Proposed AFQT score, with NO replaced by MK.

TABLE 4

QUALIFYING RATES ON ALTERNATIVE AFQTS
ARMY APPLICANTS

Group	Number of cases[a]	Percent qualifying on AFQT at percentile score: 21 Oper[b]	21 Adj[c]	21 Prop[d]	31 Oper[b]	31 Adj[c]	31 Prop[d]	50 Oper[b]	50 Adj[c]	50 Prop[d]
Males										
Total	25219	84.8	82.1	85.3	68.2	64.8	69.8	40.8	38.0	41.2
White	16519	92.0	90.1	92.7	80.3	77.2	81.8	53.0	49.6	53.4
Black	6105	70.0	65.7	70.2	42.5	38.3	43.8	14.4	12.8	14.7
Other	2595	73.2	69.8	73.9	51.4	48.7	54.2	25.2	22.9	25.7
Females										
Total	4446	89.5	87.0	88.2	74.8	71.1	71.5	43.5	40.5	39.3
White	2428	95.1	93.5	94.8	86.7	83.9	84.8	60.3	56.8	55.4
Black	1637	83.3	79.5	80.0	59.5	54.5	54.3	21.5	19.2	18.0
Other	381	80.1	77.7	81.6	64.3	60.6	61.2	31.0	27.8	27.6

---

a. Applicants for active duty tested October and November 1984.
b. Operational AFQT score, includes inflated NO scores.
c. Adjusted AFQT score, includes deflated NO scores.
d. Proposed AFQT score, with NO replaced by MK.

TABLE 5

QUALIFYING RATES ON ALTERNATIVE AFQTS
NAVY APPLICANTS

Group	Number of cases[a]	Percent qualifying on AFQT at percentile score: 21 Oper[b]	21 Adj[c]	21 Prop[d]	31 Oper[b]	31 Adj[c]	31 Prop[d]	50 Oper[b]	50 Adj[c]	50 Prop[d]
Males										
Total	12154	92.2	90.5	92.7	79.6	76.6	81.0	52.9	49.9	53.2
White	9227	95.7	94.5	96.2	86.3	83.6	87.5	61.3	58.1	61.5
Black	1892	80.1	75.9	80.0	55.3	51.4	56.3	22.5	20.0	22.6
Other	1035	83.7	81.1	83.9	64.5	61.2	68.2	33.4	31.4	35.8
Females										
Total	1956	95.2	94.1	94.0	83.7	80.8	82.1	56.0	52.9	51.3
White	1377	97.9	97.2	97.3	90.6	88.5	89.8	67.6	64.6	62.3
Black	417	89.0	86.8	84.9	65.2	60.0	61.4	25.2	21.3	20.4
Other	162	88.3	85.8	89.5	72.2	68.5	69.1	37.0	34.6	37.0

---

a. Applicants for active duty tested October and November 1984.
b. Operational AFQT score, includes inflated NO scores.
c. Adjusted AFQT score, includes deflated NO scores.
d. Proposed AFQT score, with NO replaced by MK.

TABLE 6

QUALIFYING RATES ON ALTERNATIVE AFQTS
AIR FORCE APPLICANTS

Group	Number of cases[a]	Percent qualifying on AFQT at percentile score: 21 Oper[b]	21 Adj[c]	21 Prop[d]	31 Oper[b]	31 Adj[c]	31 Prop[d]	50 Oper[b]	50 Adj[c]	50 Prop[d]
Males										
Total	8877	93.7	92.2	94.2	83.4	80.6	84.5	56.2	52.9	57.2
White	6426	97.0	96.1	97.4	89.5	87.1	90.4	64.9	61.4	65.2
Black	1732	83.3	79.8	84.4	64.2	60.0	65.0	29.9	27.3	32.4
Other	719	88.5	86.8	89.6	74.7	71.4	79.0	41.5	38.7	45.2
Females										
Total	2741	93.9	92.5	92.8	80.6	77.7	79.3	50.9	47.3	46.4
White	1799	97.1	96.4	96.7	88.7	86.1	87.7	61.9	58.3	56.9
Black	768	86.9	83.7	84.0	62.9	59.1	61.2	26.6	22.9	23.7
Other	174	92.0	89.7	92.0	74.7	73.0	73.0	44.3	40.2	39.1

---

a. Applicants for active duty tested October and November 1984.
b. Operational AFQT score, includes inflated NO scores.
c. Adjusted AFQT score, includes deflated NO scores.
d. Proposed AFQT score, with NO replaced by MK.

TABLE 7

QUALIFYING RATES ON ALTERNATIVE AFQTS
MARINE CORPS APPLICANTS

Group	Number of cases[a]	Percent qualifying on AFQT at percentile score: 21 Oper[b]	21 Adj[c]	21 Prop[d]	31 Oper[b]	31 Adj[c]	31 Prop[d]	50 Oper[b]	50 Adj[c]	50 Prop[d]
Males										
Total	6173	90.9	89.1	90.6	76.4	73.6	75.7	46.2	43.3	44.8
White	3977	95.3	94.2	95.5	85.2	82.8	85.0	58.0	54.8	56.1
Black	1531	80.5	76.9	79.0	56.1	52.4	54.5	20.6	18.8	20.4
Other	665	88.6	86.6	87.8	70.7	67.7	68.9	34.6	31.4	33.8
Females										
Total	529	96.2	94.7	94.7	86.6	83.2	83.2	58.0	54.6	52.7
White	310	98.4	97.1	98.7	93.2	91.3	90.3	70.7	68.1	65.5
Black	166	93.4	91.0	88.0	73.5	69.3	69.9	40.0	33.7	33.1
Other	53	92.5	92.5	92.5	88.7	79.3	83.0	47.2	41.5	39.6

---

a. Applicants for active duty tested October and November 1984.
b. Operational AFQT score, includes inflated NO scores.
c. Adjusted AFQT score, includes deflated NO scores.
d. Proposed AFQT score, with NO replaced by MK.

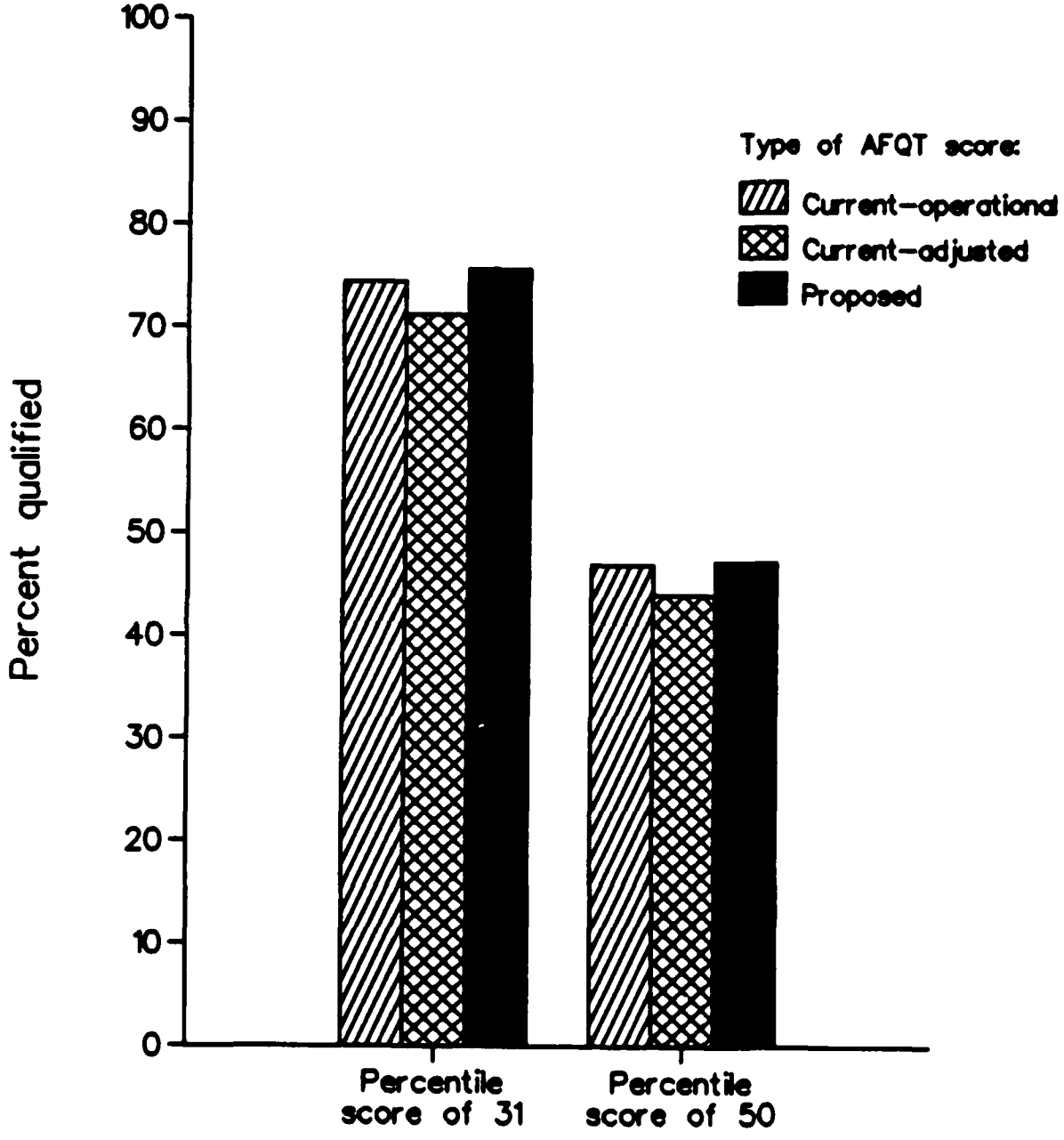


FIG. 18: QUALIFICATION RATES FOR MALES

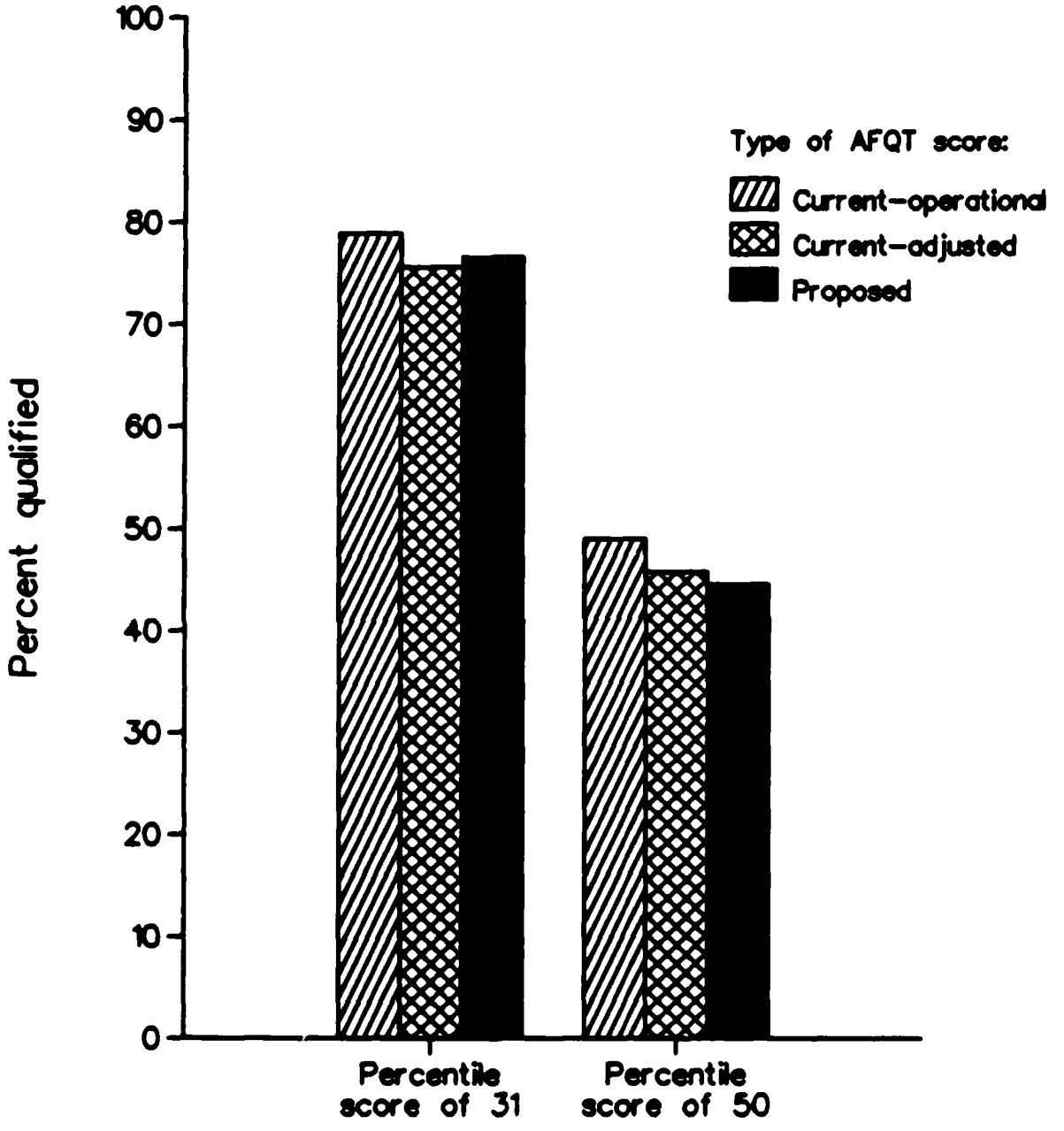


FIG. 19: QUALIFICATION RATES FOR FEMALES

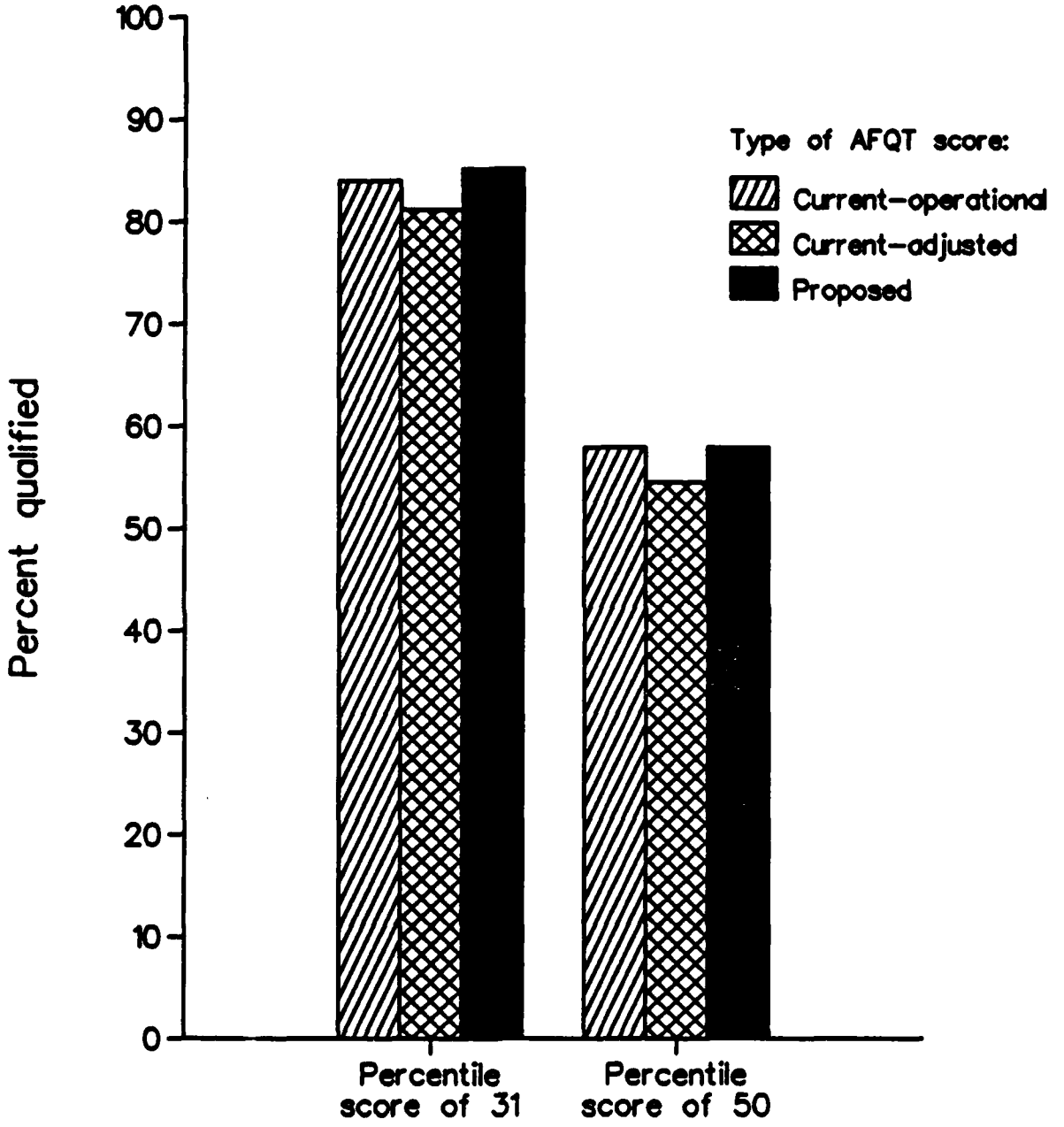


FIG. 20: QUALIFICATION RATES FOR WHITE MALES

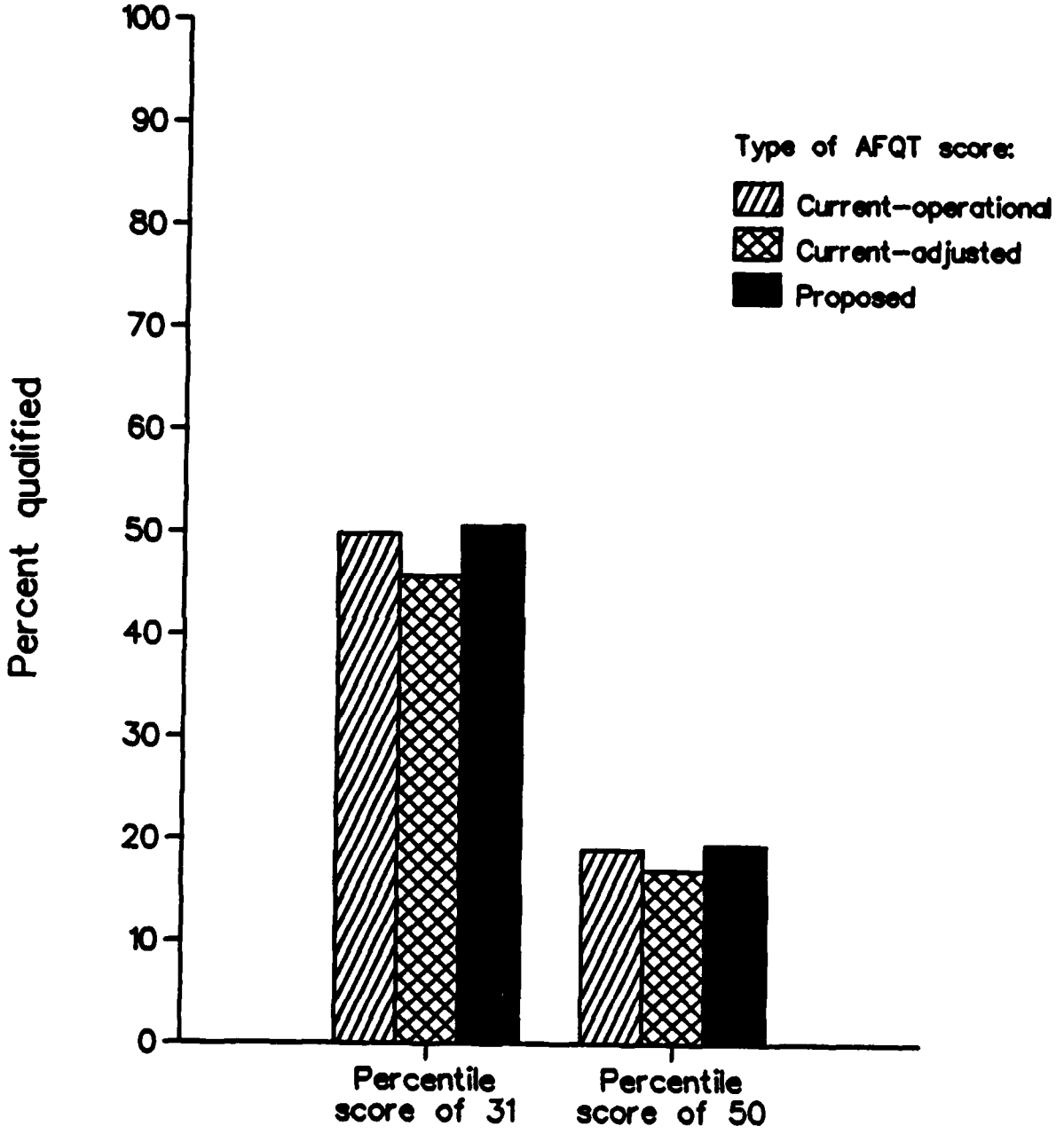


FIG. 21 : QUALIFICATION RATES FOR BLACK MALES

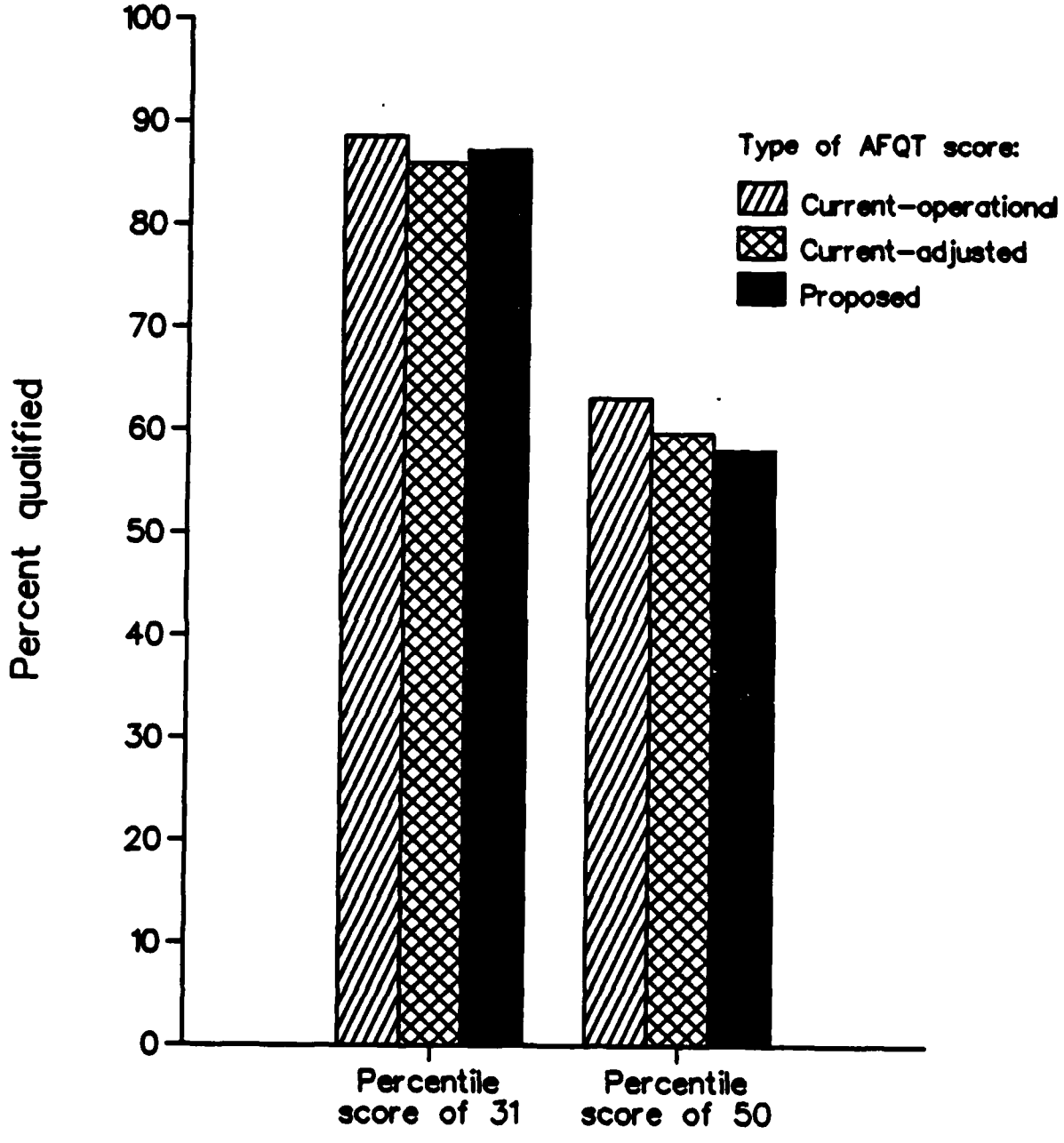


FIG. 22: QUALIFICATION RATES FOR WHITE FEMALES

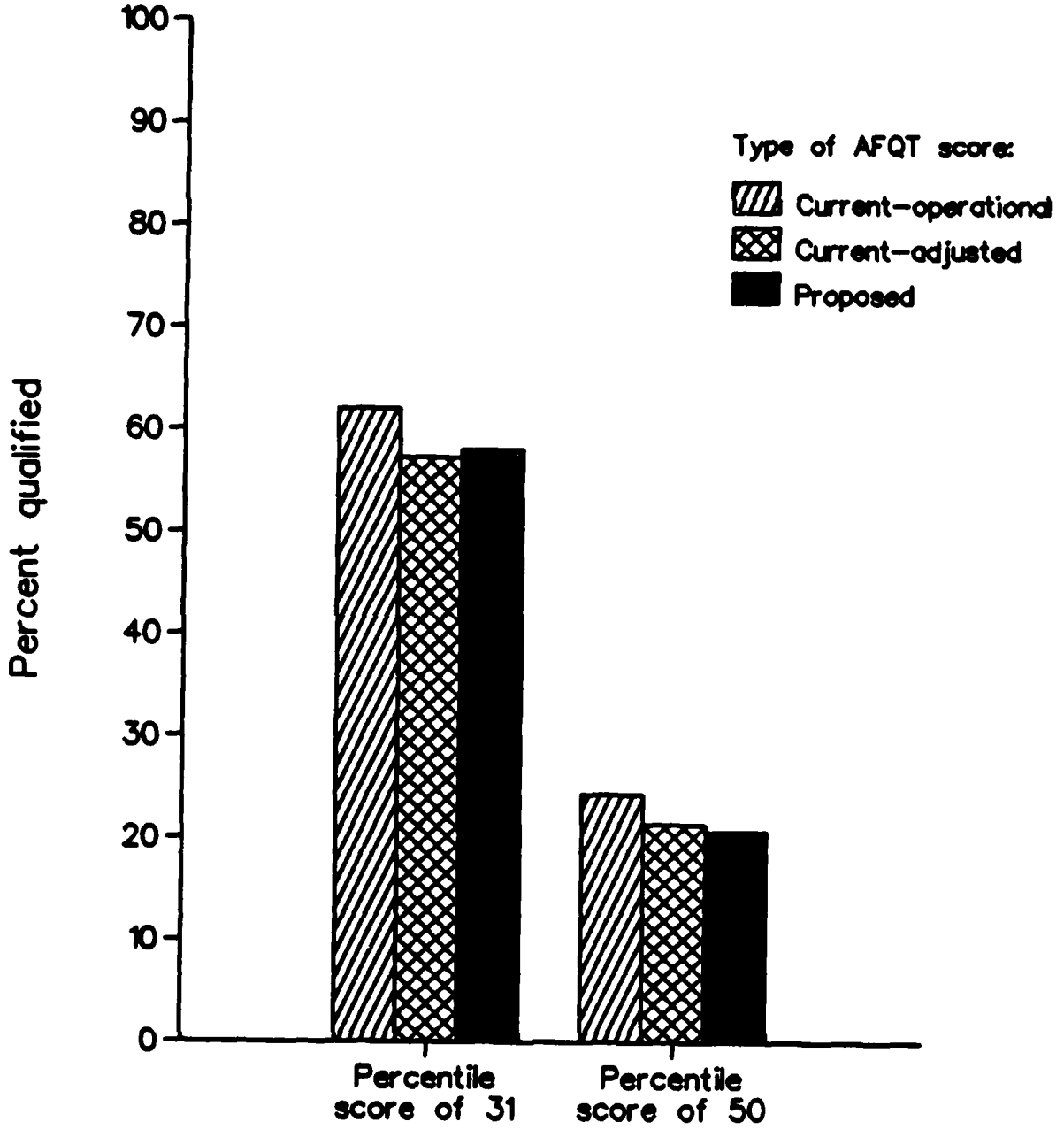


FIG. 23: QUALIFICATION RATES FOR BLACK FEMALES

Comparisons among the three sets of AFQT scores is illuminating with respect to the inflationary effect of the NO scores and the adverse impact of the proposed AFQT on females or minorities. The differences between the qualification rates based on operational and adjusted AFQT scores, both of which contain NO, reflect the effects of inflated NO scores. The differences between the rates based on the adjusted and proposed AFQT scores are the most accurate measures of the adverse impact of the proposed AFQT on females or minorities. The differences between the current operational and proposed AFQT are the changes that would be observed for current applicants if the proposed AFQT were adopted for operational use. Although the last differences are misleading because of the inflated NO scores, they need to be addressed because they are what manpower and personnel managers would evaluate if they did not know of the inflated NO scores.

Operational Versus Adjusted AFQT Scores

The differences in qualification rates between the operational and adjusted AFQT scores range up to over 10 percent of those qualified, with the differences largest for blacks. For example, 59.5 percent of the black female Army applicants were qualified at an operational AFQT score of 31; based on the adjusted AFQT scores, the rate drops to 54.5 percent, which is 5 percentage points or 8 percent of the operational rate. At an AFQT score of 50, the rate for Army black females drops from 21.5 to 19.2 percent, or about 10 percent of the operational rate. The differences for black males tend to be somewhat smaller than for black females. For whites and the "other" group, the differences tend to be smaller, both absolutely and as a percentage of the operational rates.

Considering the large numbers of applicants tested, thousands of people would have their classification lowered if the inflation of the NO scores were removed. During the 2-month period of the 1984 IOT&E (October and November 1984), 53,017 males applying for active duty were tested with forms 11, 12B, and 13; about an additional 11,000 were tested with forms 13C and 12A. The difference in qualification rate is 3.2 percentage points at percentile scores of 31 and 50 (for example, 74.4 percent qualifying on the operational scores and 71.2 percent on the adjusted scores at a percentile scores of 31). In this period, about 2,000 fewer males would have qualified for enlistment at a percentile score of 31 or for a bonus at a percentile score of 50. Thus, currently over 10,000 males each year are classified as qualified at these levels of AFQT categories but would be unqualified if the AFQT were accurately calibrated to the 1980 Youth Population.

Adjusted Versus Proposed AFQT Scores

The differences in qualification rates between the adjusted and proposed AFQT tend to be small. More males would qualify under the proposed AFQT at all three AFQT scores. For females, the differences are small and inconsistent at the lower AFQT scores, 21 and 31. At the median, 50, more females would qualify under the current AFQT. The differences are small, 1 or 2 percentage points, relative to the amount of inflation in the AFQT. The proposed AFQT would have virtually no impact on the true qualification rates of minority females (blacks and others); the largest adverse impact would be on white females. For example, 59.7 percent of the total DOD white females would qualify at a percentile score of 50 on the adjusted AFQT while 58.1 percent would qualify on the proposed AFQT. This difference of 1.6 percentage points would result in about 800 fewer white females each year classified into AFQT category IIIA (percentile scores of 50 through 64).[1]

Operational Versus Proposed AFQT Scores

The observed drop in qualification rates between the AFQT scores of record, based on the inflated AFQT scores, and the proposed AFQT would be large for females. The qualification rates for females would drop about 10 percent. About 80,000 females applied for active duty in FY 1985, and about 50,000 were classified into AFQT category IIIA. Under the proposed AFQT, about 5,000 fewer females would be placed into this category. These differences hold for the three female racial groups.

For males, the differences are nominal. There is a small tendency for more males to be qualified under the proposed AFQT. The net effect of adopting the proposed AFQT would be that the percentage of qualified males would appear to increase slightly compared to the operational scores, but in fact would increase noticeably, by over 5 percent, at percentile scores of 31 and 50 if the AFQTs were accurately calibrated to the 1980 Youth Population.

1. About 8,400 white females (5,978 tested with forms 11, 12B, and 13 plus 2,383 tested with forms 12A and 13C) applied for active duty during the 1984 IOT&E. For a full year, about 51,000 would apply, and 1.6 percent of 51,000 is 816.

Qualification Rates in the 1980 Youth Population

The qualification rates on each set of AFQT scores were also computed for the 1980 Youth Population. These qualification rates are used primarily to evaluate the effects of inflated NO scores for examinees who have not been prescreened prior to testing with the ASVAB. Most applicants for enlistment have been screened by recruiters on the Enlistment Screening Test, and people clearly unqualified for enlistment are not forwarded to examining stations. Because of the prescreening, qualification rates for current applicants are higher than they would be for representative groups of examinees.

The qualification rates for the 1980 Youth Population are the rates that would be expected if the people tested with the ASVAB were representative of the current youth population, as in times of mobilization. The three AFQT scores are as follows: current accurate, which corresponds to the current AFQT score scale in the 1980 Youth Population; current inflated, with the NO scores inflated by the amounts found for the 1984 IOT&E group; and proposed, in which MK replaces NO.

During times of mobilization, the NO scores probably would not be inflated to the same degree as they are currently for applicants. During mobilization large numbers would be drafted, and registrants for induction do not have the same incentive to score well as do applicants for enlistment; also they are not as likely to be aware of the test-taking strategies.

The qualification rates of AFQT percentile scores 21, 31, and 50 are shown in table 8 for each sex and racial group (whites, blacks, and others). The rates for males and females are shown in figures 24 and 25, respectively.

DISCUSSION

The initial focus of this analysis was to evaluate the impact of the proposed AFQT, in which MK would replace NO, on qualification rates of applicants compared to the current AFQT. The focus shifted, however, as the magnitude of the inflated NO scores become apparent—up to 8 points for males and 7 points for females, with the AFQT scores consequently inflated by up to 3 percentile score points. This inflation means that the current AFQT scores are not accurately calibrated to the 1980 Youth Population and that applicants are systematically misclassified as having higher aptitudes than they do in fact have.

TABLE 8

QUALIFYING RATES ON ALTERNATIVE AFQTS
1980 YOUTH POPULATION

Group[a]	Percent qualifying on AFQT at percentile score								
	21			31			50		
	Infl[b]	Crct[c]	Prop[d]	Infl[b]	Crct[c]	Prop[d]	Infl[b]	Crct[c]	Prop[d]
Males									
Total	80.4	78.6	80.3	71.0	69.3	70.6	54.0	52.0	53.4
White	87.6	86.3	87.8	79.7	78.0	79.2	62.4	60.5	61.6
Black	45.7	42.1	44.6	29.3	27.8	30.4	14.7	12.7	15.3
Other	45.7	57.8	58.9	47.3	45.7	46.3	29.2	26.5	28.9
Females									
Total	82.2	80.8	80.6	72.9	69.7	69.4	51.8	49.4	48.1
White	89.8	88.8	88.6	82.5	79.2	78.4	60.7	58.1	56.4
Black	49.6	45.7	46.7	31.9	29.2	31.8	14.0	12.4	13.0
Other	57.2	54.6	53.2	41.1	38.2	37.5	21.1	19.4	18.6

---

a. In the 1980 Youth Population, "Other" includes only Hispanics; "White" includes Asians and other minorities.
b. Inflated AFQT score, includes inflated NO scores.
c. Correct AFQT score, includes correct NO scores.
d. Proposed AFQT score, with NO replaced by MK.

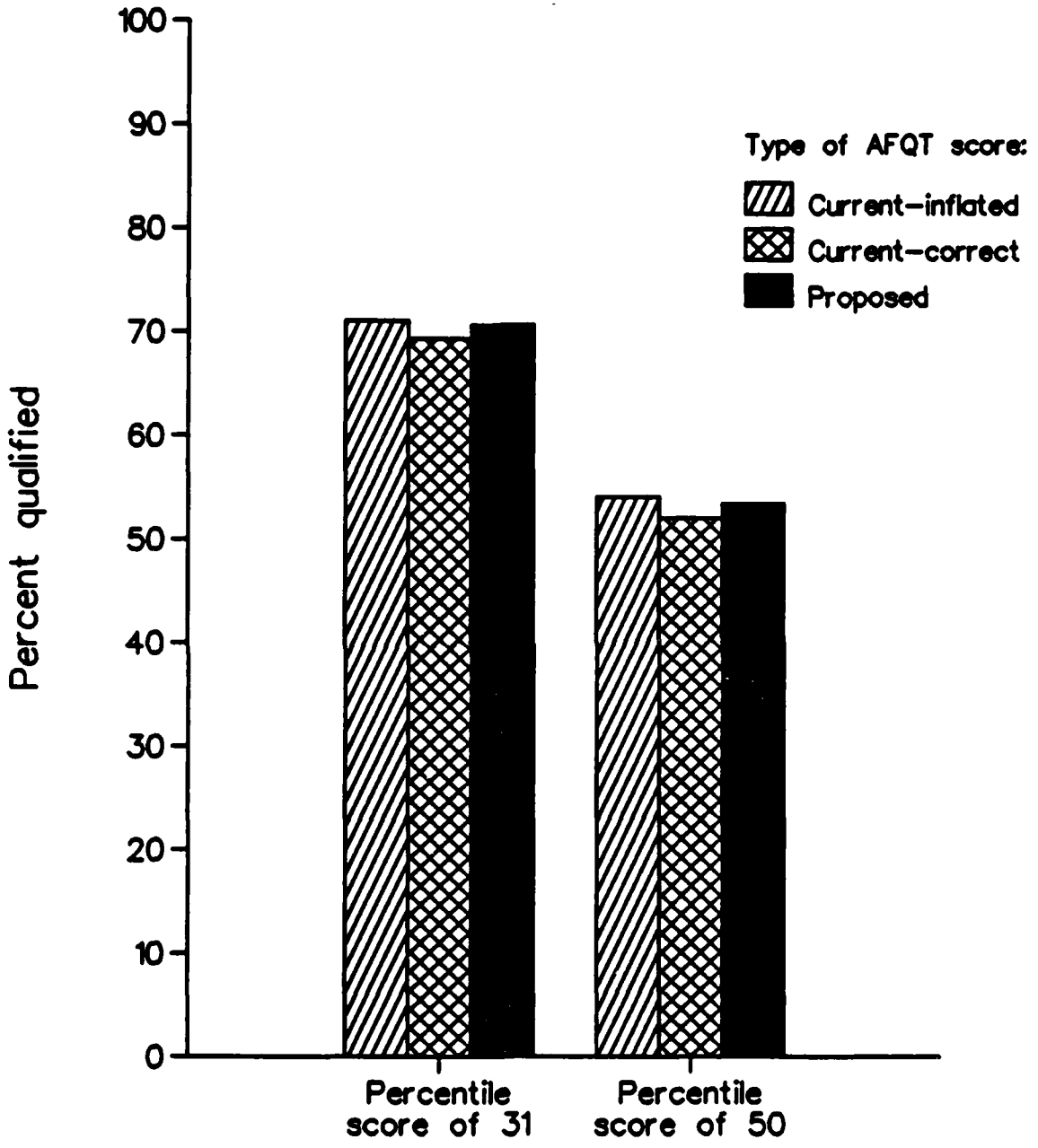


FIG. 24: QUALIFICATION RATES FOR MALES IN THE 1980 YOUTH POPULATION

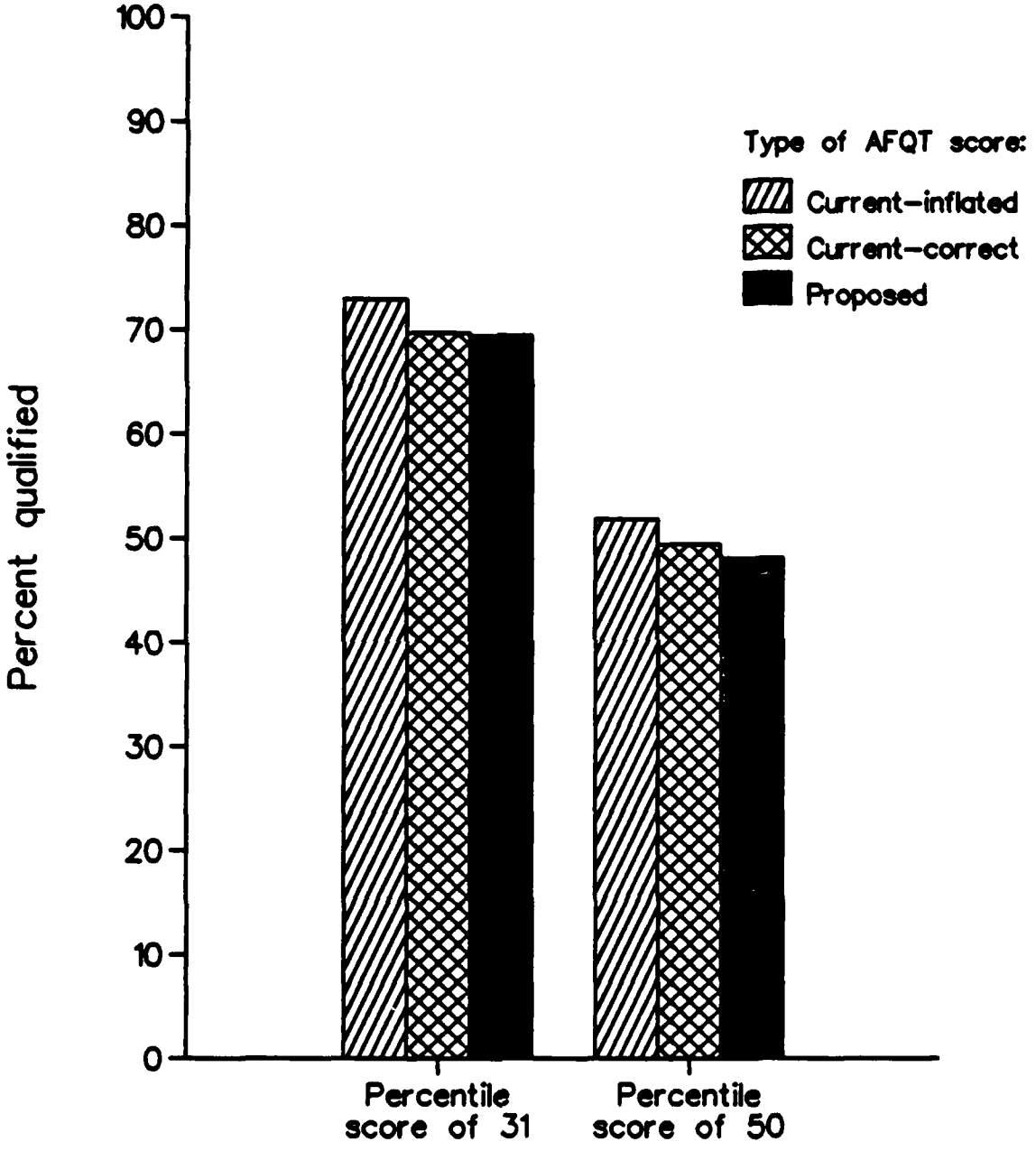


FIG. 25: QUALIFICATION RATES FOR FEMALES IN THE 1980 YOUTH POPULATION

The accuracy of the ASVAB scores is the fundamental requirement for using them in making personnel decisions. In 1980 DOD went to much effort and expense to develop a new reference population for the ASVAB. The new score scale and test norms were introduced on 1 October 1984, along with new forms of the ASVAB (ASVAB 11/12/13).

The Joint Services Selection and Classification Working Group diligently monitors the ASVAB scores to maintain their accuracy. The primary vehicle for evaluating their accuracy is through an Initial Operational Test and Evaluation (IOT&E) conducted immediately after a new version of the ASVAB is introduced. Score distributions are also examined each month to look for anomalies.

An adjustment to the calibration of ASVAB 11/12/13 was made on 1 July 1986 because the AFQT scores were found to be approximately 2 percentile score points too low compared to the 1980 Youth Population. DOD personnel managers decided that an error of such magnitude in the ASVAB score scale is intolerable, and the adjustment was publicly acknowledged.

An error of similar magnitude, but of opposite direction, is now known to be present in the AFQT scores. The source of the error, as has been true of other calibration errors since 1980, is the NO subtest.

The July 1986 adjustment was required because the AFQT scores were too low by approximately 2 percentile score points. The inflation of NO scores found in the analysis raises them by 1 to 3 percentile score points. Ironically, the two effects generally balanced each other out between October 1984 and July 1986; classification decisions based on the AFQT during that period were essentially accurate. After July 1986, the deflation effect was removed, and only the inflation effect remains. Unknown to the Working Group and DOD personnel managers, prior to July 1986 the NO subtest had inadvertently maintained reasonable equity of classification decisions. The evidence presented in this report indicates that equity can be restored and maintained only by deleting NO from the AFQT.

The question then arises whether NO has any place in the ASVAB. It is used by all services, except the Marine Corps, in aptitude composites, which are combinations of ASVAB subtests used to determine the qualification of recruits for assignment to occupational specialties. The primary requirement of aptitude composites is that they have validity for predicting performance in the various specialties open to recruits. In the validation studies conducted by

the services in the early 1980s, NO was found to have sufficient validity to warrant its inclusion in some aptitude composites. The ASVAB test scores used in these validation studies, however, were for recruits that started their occupational specialty training courses in fiscal years 1981 and 1982. During that period, the NO scores may have been relatively accurate, and the validity of NO may not have been degraded by the test-taking strategies. At the present time, NO scores are inflated, and the predictive validity of the inflated scores is not known. Validation studies using the inflated NO scores are required to evaluate the usefulness of the inflated NO scores in aptitude composites.

The same argument about retaining NO in the ASVAB applies to Coding Speed (CS), the other speeded test in the battery. Even though CS is not part of the AFQT, the CS scores are inflated to about the same magnitude as NO. Analysis of the CS scores is presented in appendix A. The validity of CS in the validation studies conducted in the early 1980s was satisfactory. However, the predictive validity of the inflated CS scores also needs to be evaluated.

QUALIFICATION RATES

The impetus for this analysis was to evaluate the differences in qualification rates for the current and proposed AFQT. Because females tend to do better on speeded tests than males, the change from NO to MK in the AFQT was expected to have an adverse impact on females. The analysis focused on qualification rates at points where enlistment decisions tend to be made, AFQT percentile scores of 21 and 31, and where decisions about bonuses and guaranteed assignments tend to be made, AFQT percentile score of 50. As expected, the proposed AFQT would result in fewer females qualifying. The differences in qualification rates between the proposed AFQT and adjusted AFQT scores are noticeably smaller than between the proposed AFQT and operational AFQT scores (generally from 2 to 4 percentage points lower). Examination of qualification rates in extant AFQT score distributions would include the inflated NO scores and therefore overestimate the adverse impact on females.

Females do have lower qualification rates on the proposed AFQT compared to the adjusted AFQT. These are true differences, and they are expected to persist. The recommendation to change the AFQT by replacing NO with MK coupled with inherent defects of NO creates a dilemma for the DOD testing community and manpower managers. Retaining NO would perpetuate known errors in the scores; replacing NO would have an adverse

impact on a protected minority, females. The resolution seems to rest on which of the consequences is less undesirable. Perpetuating a known error is hardly justifiable, which leaves coping with the adverse impact as the viable alternative. The drop in qualification rates would be relatively small, especially considering that recruiting enough females tends to be less of a problem for the services than recruiting enough males.

ACCURACY OF ADJUSTING THE NO SCORES

The adjustment to the NO scores that takes into account inflation arising from test-taking strategies is accurate on average; the qualification rates computed for the adjusted AFQT scores should reasonably well represent the percentage of applicants who would be qualified on the AFQT when it is accurately calibrated to the 1980 Youth Population. For purposes of making classification decisions about individuals, however, the adjustment may or may not be accurate. The adjustment reflects an average change in NO scores. For people who used the test-taking strategies to the full extent, the adjustment is too small. Conversely for people who are unaware of the test-taking strategies, or choose to ignore some or all of them, the adjustment is too large. On average, equity can be attained through the adjustment, but for individual decisions, large classification errors can be made.

The average inflation effect was found to be higher for people who want to serve on active duty than for those who want to join a reserve component. Also, the effect was larger for people retesting than for those taking their initial test. Finding a single correct adjustment would be impossible, and having multiple adjustments would be horrendously complicated. Adjusting the NO scores for inflation is virtually impossible.

The argument about the accuracy of adjustments for computing qualification rates versus making decisions about individuals also applies to the previous adjustments to the NO-scores to correct for the effects of answer sheet design and type font—although to a lesser extent. Some individuals may have been affected more than others by filling in circles versus rectangles on the answer sheet or by the size of the type. To the extent that an individual deviates from the average, the adjustment is not accurate. The adjusted NO scores, and resulting AFQT scores, are far superior to the unadjusted ones, but still less than perfect.

The best, and only sure way, to calibrate the AFQT accurately to the 1980 Youth Population is to delete NO. As long as NO remains in the AFQT, equity in classfication decisions remains doubtful.

RECOMMENDATIONS

Recommendations from the analysis are the following:

- Delete NO from the AFQT and replace it with MK.
- Evaluate the predictive validity of the inflated NO scores to determine whether it should be retained in the ASVAB.

The NO subtest, and the CS subtest as well, is inherently flawed for use in an operational testing program where examinees have prior information about the types of test content. Examinees can legitimately prepare themselves to improve their speeded test scores without improving the underlying ability that would also improve performance in their occupational specialties. Continued use of NO in the AFQT would negate the efforts to maintain the accuracy of the ASVAB score scale.

REFERENCES

[1] CNA, Memorandum 83-3102, *The Appropriateness for Military Applicants of the ASVAB Subtests and Score Scale in the New 1980 Reference Population*, by William H. Sims and Milton H. Maier, Unclassified, Jun 1983

[2] CNA, Memorandum 82-3118, *Constructing an ASVAB Score Scale in the 1980 Reference Population*, by Milton H. Maier and William H. Sims, Unclassified, Aug 1982

[3] Air Force Human Resources Laboratory, *The 1980 Youth Population: An Investigation of Speeded Subtests*, by James A. Earles, Toni Giulliano Wegner, Malcolm J. Ree, and Lonnie D. Valentine, Jr., Unclassified, Aug 1983

[4] Army Research Institute, *The Effects of Practice on the Armed Services Vocational Aptitude Battery*, by Bruce K. McCormick, Robert S. Kennedy, and Marshall B. Jones, Unclassified, Dec 1983

[5] Army Research Institute, *A Study of the Reliability of Scores for Fiscal Year 1981 Army Applicants*, Briefing to the Joint Services Selection and Classification Working Group, Minutes, 14 Sep 1982

APPENDIX A

COMPUTING THE AMOUNT OF INFLATION IN THE NO AND CS RAW SCORES

APPENDIX A

COMPUTING THE AMOUNT OF INFLATION IN THE NO AND CS RAW SCORES

The amount of inflation in the Numerical Operations (NO) and Coding Speed (CS) raw scores was computed by comparing their distribution in the 1980 Youth Population and the 1984 IOT&E. NO and CS scores in the 1984 IOT&E were equated to those in the 1980 Youth Population. Raw scores (number of items correct) that had the same cumulative frequency in the two groups were set equal, or equated, to each other. The amount of inflation is the difference between raw scores that have the same cumulative frequencies in the two groups.

Prior to computing the cumulative frequencies of the NO and CS, the two groups had to be weighted to have the same distribution of ability. Cumulative frequencies can be compared only if the groups are equivalent in relevant abilities, which in this case is a measure of general ability. The Health, Social, and Technology (HST) occupational composite, composed of Verbal (VE), Arithmetic Reasoning (AR) and Mechanical Comprehension (MC), was used to weight the two groups.

The HST scores were grouped into intervals of 10 points. Weights were assigned to the people in each interval such that the weighted frequencies were the same for the two groups. With only two groups, one distribution ordinarily is transformed to equal the other. This weighting poses problems for equating. If the 1984 IOT&E were weighted to have the same distribution of HST scores as the 1980 Youth Population, the NO and CS scores in the 1984 IOT&E would be transformed only to the extent that they are correlated with HST. If they were uncorrelated, their distribution would not change, and the weighting would not make them equivalent with respect to NO and CS. A preferred solution is to weight the two groups of interest to have the same distribution as a third group. Examinees in the 1980 IOT&E was used as the third group.

Three separate equatings for males of the NO and CS scores were computed to evaluate the effects of weighting on the outcome. Each group in turn was used as the reference – the 1980 Youth Population, the 1984 IOT&E, and the 1980 IOT&E – and the other two groups were weighted to have the same distributions of HST scores. The weights are shown in table A-1.

TABLE A-1

STRATIFICATION WEIGHTS FOR ACTIVE DUTY APPLICANTS

Panel A: Uses 1980 IOT&E as Reference

Sum of subtest standard scores	Males	
	1980 Youth	1984 IOT&E
110	1.224	4.151
111-120	1.634	2.129
121-130	1.515	1.238
131-140	1.550	0.9[illegible]
141-150	1.305	0.6[illegible]8
151-160	1.112	0.7[illegible]
161-170	0.820	0.74[illegible]
171-180	0.667	0.782
181-190	0.518	0.763
190	0.419	1.000

Panel B: Uses 1980 Youth Population as Reference

Sum of subtest standard scores	Males		Females
	1980 IOT&E	1984 IOT&E	1984 IOT&E
110	0.817	3.392	3.121
111-120	0.612	1.305	1.123
121-130	0.660	0.818	0.734
131-140	0.645	0.634	0.707
141-150	0.766	0.642	0.757
151-160	0.899	0.707	0.909
161-170	1.219	0.903	1.065
171-180	1.500	1.173	1.589
181-190	1.930	1.474	1.733
190	2.384	2.382	1.733

Panel C: Uses ASVAB Form 13C from 1984 IOT&E as Reference

Sum of subtest standard scores	Males		Females
	1980 IOT&E	1980 Youth	1980 Youth
110	0.241	0.295	0.320
111-120	0.470	0.766	0.890
121-130	0.808	1.222	1.362
131-140	0.017	1.578	1.414
141-150	1.193	1.557	1.321
151-160	1.271	1.414	1.100
161-170	1.350	1.108	0.939
171-180	1.279	0.853	0.629
181-190	1.311	0.678	0.577
190	1.000	0.420	0.577

A separate equating was performed for females. One reason is that the 1980 IOT&E did not contain any females. Another is that the test-taking strategies may be different for females, and the amount of inflation could be lower or higher than for males. The weights for females are also shown in table A-1.

The NO scores in 1984 IOT&E equivalent to scores in the 1980 Youth Population are shown in table A-2 for males and table A-3 for females. Three sets of equivalent scores are shown for males, and two for females. Each set has a different group as the reference to which the others were weighted. The similarity of the results using the three different reference groups suggests that the equatings are reliable.

For males in the 1984 IOT&E, the equivalent NO scores tended to be 1 point higher when the 1980 Youth Population was the reference than when the 1984 IOT&E was used as the reference, up to a score of 32. For most NO scores up to 28, use of the 1980 Youth Population and the 1980 IOT&E as the references resulted in identical equivalent scores. Above a score of 28, the equivalent scores were identical, and use of the 1980 IOT&E as the reference resulted in equivalent scores 1 point higher than when the other two groups were used as the reference. The amount of inflation, shown in table 4 of the main text, was measured by using the 1980 IOT&E as the reference.

For females in the 1984 IOT&E, the equivalent NO scores, up to a score of 23, were lower when the 1980 Youth Population was used as the reference. This result is opposite to that for males. Above a NO score of 23, the two reference groups tended to result in identical equivalent scores. The final amount of inflation of NO and scores below 24 for females, as shown in table 4 of the main text, was based primarily on the results using the 1980 Youth Population as the reference, which minimized the inflation.

Equating of the CS scores is shown in table A-4. The amount of inflation for CS is comparable to that for NO. The maximum inflation for NO was 8 points, which is about three-fourths of a standard deviation. For CS the maximum inflation was 14 points, which is also about three-fourths of a standard deviation.

Inflation of the NO scores was expected because of the emphasis among applicants and recruiters on attaining high AFQT scores, which contains the NO subtest. CS is not part of the AFQT, and therefore is less subject to explicit preparation to improve test scores. In fact, CS is part of the pseudo-AFQT, a composite of ASVAB subtests used to help identify cheating on the

AFQT. The inflated CS scores inflate the pseudo-AFQT scores, which compensates for the inflated NO scores arising from test-taking strategies. The pseudo-AFQT, therefore, remains a valid check on cheating, which is expectd to affect the Word Knowledge and Arithmetic Reasoning scores more than NO. No change to the pseudo-AFQT is required because of the inflated NO and CS scores.

TABLE A-2

EQUATING OF NO SCORES FOR MALES

NO score in 1980 Youth Population	Equivalent NO score in 1984 IOT&E group		
	a	b	c
	A	B	C
1	9	9	9
2	10	10	10
3	11	11	10
4	12	12	11
5	13	14	12
6	14	15	13
7	14	15	13
8	15	16	14
9	16	17	15
10	17	18	17
11	18	18	18
12	19	19	18
13	19	20	19
14	20	20	19
15	21	21	20
16	22	22	21
17	23	23	22
18	24	24	23
19	24	24	23
20	25	25	24
21	25	25	25
22	26	26	26
23	27	27	26
24	28	28	27
25	29	29	28
26	30	30	29
27	31	31	30
28	32	32	31
29	33	32	32
30	34	33	33
31	35	34	34
32	35	34	34
33	36	35	35
34	37	36	36
35	38	37	37
36	39	37	37
37	40	38	38
38	41	39	39
39	42	41	41
40	42	42	42

TABLE A-2 (Con't.)

NO score in 1980 Youth Population	Equivalent NO score in 1984 IOT&E group		
	a	b	c
	A	B	C
41	43	42	43
42	44	43	43
43	45	44	44
44	45	45	45
45	46	46	46
46	47	46	46
47	47	47	47
48	48	47	48
49	49	48	49
50	50	50	50

a. Uses 1980 IOT&E as reference.
b. Uses 1980 Youth Population as reference.
c. Uses 1984 IOT&E as reference.

TABLE A-3

EQUATING OF NO SCORES FOR FEMALES

NO score in 1980 Youth Population	Equivalent NO score in 1984 IOT&E group	
	a	c
	A	B
1	8	2
2	10	7
3	11	10
4	12	11
5	13	12
6	14	12
7	15	13
8	16	14
9	17	15
10	18	16
11	19	17
12	20	19
13	21	20
14	22	21
15	22	22
16	23	22
17	23	23
18	24	23
19	25	24
20	25	24
21	26	25
22	27	26
23	28	27
24	28	28
25	29	29
26	30	30
27	32	31
28	33	32
29	34	33
30	35	34
31	35	35
32	36	36
33	37	36
34	38	37
35	38	38
36	39	39
37	40	40
38	41	41
39	43	43
40	44	44

TABLE A-3 (Con't.)

NO score in 1980 Youth Population	Equivalent NO score in 1984 IOT&E group	
	a A	b B
41	45	45
42	45	46
43	46	46
44	47	47
45	47	47
46	47	48
47	48	48
48	48	48
49	49	49
50	50	50

a. Uses 1980 Youth Population as reference.
b. Uses 1984 IOT&E as reference.

TABLE A-4

EQUIVALENT CS SCORES FOR THE 1980 YOUTH POPULATION AND THE 1984 IOT&E EXAMINEES

CS score in the 1980 Youth Population	1984 IOT&E examinees	
	Males	Females
1	1	1-6
2	2-10	7-10
3	11	11-12
4	13	13-14
5	14	15-16
6	15	17-19
7	16	20-21
8	17	22
9	18	23
10	19	24
11	20	25
12	22	26
13	24	27
14	26	28
15	27	29
16	28	30
17	29	31
18	30	32
19	30	33
20	31	34
21	32	34
22	32	35
23	33	36
24	34	36
25	34	37
26	35	38
27	36	28
28	36	39
29	37	39
30	38	40
31	39	40
32	39	41
33	40	41
34	41	42
35	42	42
36	42	43
37	43	43
38	44	44
39	45	45
40	45	46

TABLE A-4 (Con't.)

CS score in the 1980 Youth Population	1984 IOT&E examinees	
	Males	Females
41	46	46
42	47	47
43	48	48
44	48	49
45	49	49
46	50	50
47	51	51
48	52	52
49	53	52
50	53	53
51	54	54
52	54	54
53	55	55
54	56	56
55	57	57
56	58	57
57	59	58
58	60	59
59	61	61
60	62	62
61	63	63
62	64	64
63	65	65
64	66	66
65	66	67
66	67	68
67	67	69
68	68	69
69	70	70
70	71	71
71	72	72
72	74	73
73	75	75
74	76	76
75	77	77
76	78	78
77	79	79
78	80	80
79	81	80
80	82	81
81	82	82
82	82	82
83	83	83
84	84	84

APPENDIX B

SCALING THE PROPOSED AFQT TO THE 1980 METRIC

APPENDIX B

SCALING THE PROPOSED AFQT TO THE 1980 METRIC

The proposed AFQT, composed of VE + AR + MK, has not been officially scaled to the 1980 metric. Because all subtests in the ASVAB were administered to the 1980 Youth Population, scaling the proposed AFQT is relatively simple. The steps are as follows:

- Compute the cumulative frequency distribution (CFD) of the sum of VE, AR, and MK raw scores for form 8A in the 1980 Youth Population, and convert the sum of raw scores to percentile scores. This operation produces a score scale for the proposed AFQT.

- Compute the CFD of the proposed AFQT for form 13C, which is identical to form 8A, administered during the 1984 IOT&E, and compute the CFD of the proposed AFQT for forms 11, 12B, and 13 in the 1984 IOT&E. Note that form 12A is excluded because it is not parallel to the other forms.

- Perform equipercentile equating of the sum of raw scores for form 13C and forms 11, 12B, and 13. The conversion from sum of raw scores for forms 11, 12, and 13 to percentile scores is obtained through the equivalent sum of raw scores for form 13C (8A).

The CFD for form 8A in the 1980 Youth Population is shown in table B-1. The frequencies were smoothed by using the 3-point moving average, with weights of .25, .50 and .25. The progression of percentile scores is orderly. All crucial percentile scores (9, 10, 16, 21, 31, 63, and 93) occur naturally, except for a percentile score of 50, the median. The Joint Services Selection and Classification Working Group will need to decide which raw score to convert to a percentile score of 50. For purposes of this study, raw scores of 71 and above were said to be above a percentile score of 50.

The CFDs for forms 13C are shown in table B-2 and for forms 11, 12B, and 13 in table B-3. Raw scores that had the same cumulative frequency were set equal to each other.

The conversion from sum of subtest raw scores for the proposed AFQT to percentile scores for form 11, 12B, and 13 is shown in table B-4. This conversion table was used to calculate the qualification rates for applicants in the 1984 IOT&E reported in the main text.

TABLE B-1

CUMULATIVE DISTRIBUTION OF PROPOSED AFQT
IN THE 1980 YOUTH POPULATION

Value	Freq	Percent	Cum percent
.00	22942	.1	.1
1.00	3793	.0	.1
2.00	1064	.0	.1
3.00	925	.0	.1
4.00	985	.0	.1
6.00	1605	.0	.1
7.00	4599	.0	.1
8.00	4873	.0	.2
9.00	1121	.0	.2
11.00	6042	.0	.2
12.00	5542	.0	.2
14.00	15949	.1	.3
15.00	7086	.0	.3
16.00	11068	.0	.3
17.00	25969	.1	.4
18.00	41749	.2	.6
19.00	44356	.2	.8
20.00	53221	.2	1.0
21.00	48470	.2	1.2
22.00	93015	.4	1.6
23.00	101522	.4	2.0
24.00	100370	.4	2.3
25.00	136430	.5	2.9
26.00	123362	.5	3.4
27.00	136482	.5	3.9
28.00	150365	.6	4.5
29.00	143203	.6	5.1
30.00	155374	.6	5.7
31.00	175500	.7	6.4
32.00	202674	.8	7.2
33.00	150707	.6	7.8
34.00	183257	.7	8.5
35.00	207127	.8	9.3
36.00	152898	.6	9.9
37.00	216797	.9	10.7
38.00	195509	.8	11.5
39.00	220085	.9	12.4
40.00	205447	.8	13.2

TABLE B-1 (Con't.)

Value	Freq	Percent	Cum percent
41.00	205232	.8	14.0
42.00	229282	.9	14.9
43.00	240367	.9	15.8
44.00	222350	.9	16.7
45.00	211830	.8	17.6
46.00	233469	.9	18.5
47.00	279378	1.1	19.6
48.00	318672	1.3	20.8
49.00	264614	1.0	21.9
50.00	268869	1.1	22.9
51.00	219194	.9	23.8
52.00	338084	1.3	25.1
53.00	279645	1.1	26.2
54.00	251772	1.0	27.2
55.00	317153	1.2	28.5
56.00	388644	1.5	30.0
57.00	277724	1.1	31.1
58.00	298597	1.2	32.3
59.00	354063	1.4	33.7
60.00	301647	1.2	34.8
61.00	323671	1.3	36.1
62.00	317452	1.2	37.4
63.00	419449	1.7	39.0
64.00	372286	1.5	40.5
65.00	389240	1.5	42.0
66.00	336764	1.3	43.3
67.00	317732	1.3	44.6
68.00	385583	1.5	46.1
69.00	418759	1.6	47.8
70.00	373038	1.5	49.2
71.00	367863	1.4	50.7
72.00	342950	1.3	52.0
73.00	336281	1.3	53.3
74.00	364493	1.4	54.8
75.00	424764	1.7	56.4
76.00	323275	1.3	57.7
77.00	465650	1.8	59.6
78.00	465397	1.8	61.4
79.00	445203	1.8	63.1
80.00	376320	1.5	64.6

TABLE B-1 (Con't.)

Value	Freq	Percent	Cum percent
81.00	424941	1.7	66.3
82.00	409867	1.6	67.9
83.00	371187	1.5	69.4
84.00	366507	1.4	70.8
85.00	397010	1.6	72.4
86.00	356877	1.4	73.8
87.00	389081	1.5	75.3
88.00	388750	1.5	76.8
89.00	399212	1.6	78.4
90.00	335837	1.3	79.7
91.00	336280	1.3	81.1
92.00	317709	1.3	82.3
93.00	448672	1.8	84.1
94.00	384860	1.5	85.6
95.00	393476	1.5	87.1
96.00	465521	1.8	89.0
97.00	362582	1.4	90.4
98.00	350793	1.4	91.8
99.00	397475	1.6	93.3
100.00	421743	1.7	95.0
101.00	359683	1.4	96.4
102.00	347824	1.4	97.8
103.00	305996	1.2	99.0
104.00	188558	.7	99.7
105.00	68342	.3	100.0

TABLE B-2

CUMULATIVE FREQUENCY DISTRIBUTION OF PROPOSED AFQT FOR THE 1984 IOT&E EXAMINEES TESTED WITH ASVAB FORM 13C

Value	Frequency	Percent	Cum percent
16.00	1	0	0
17.00	2	0	0
19.00	1	0	0
21.00	8	1	1
22.00	6	0	1
23.00	8	1	2
24.00	11	1	2
25.00	19	1	4
26.00	18	1	5
27.00	24	2	7
28.00	25	2	8
29.00	40	3	11
30.00	34	2	13
31.00	41	3	16
32.00	46	3	19
33.00	46	3	22
34.00	59	4	26
35.00	69	5	30
36.00	69	5	35
37.00	75	5	40
38.00	83	6	45
39.00	87	6	51
40.00	113	8	59
41.00	117	8	67
42.00	120	8	74
43.00	104	7	81
44.00	139	9	91
45.00	134	9	100
46.00	136	9	109
47.00	174	1.2	120
48.00	205	1.4	134
49.00	193	1.3	146
50.00	181	1.2	159
51.00	217	1.4	173
52.00	255	1.7	190
53.00	246	1.6	206
54.00	261	1.7	223
55.00	258	1.7	241
56.00	261	1.7	258
57.00	299	2.0	278
58.00	277	1.8	296
59.00	296	2.0	316
60.00	254	1.7	333

TABLE B-2 (Con't.)

Value	Frequency	Percent	Cum percent
61.00	285	1.9	35.2
62.00	291	1.9	37.1
63.00	286	1.9	39.0
64.00	332	2.2	41.2
65.00	287	1.9	43.1
66.00	322	2.1	45.2
67 00	295	2.0	47.2
68.00	286	1.9	49.1
69.00	299	2.0	51.1
70.00	302	2.0	53.1
71.00	297	2.0	55.1
72.00	286	1.9	57.0
73.00	286	1.9	58.9
74.00	307	2.0	60.9
75.00	296	2.0	62.9
76.00	281	1.9	64.7
77.00	264	1.8	66.5
78.00	239	1.6	68.1
79.00	264	1.8	69.8
80.00	274	1.8	71.6
81.00	262	1.7	73.4
82.00	254	1.7	75.1
83.00	237	1.6	76.6
84.00	233	1.5	78.2
85.00	242	1.6	79.8
86.00	238	1.6	81.4
87.00	222	1.5	82.8
88.00	185	1.2	84.1
89.00	200	1.3	85.4
90.00	205	1.4	86.8
91.00	217	1.4	88.2
92.00	203	1.3	89.5
93.00	176	1.2	90.7
94.00	184	1.2	91.9
95.00	186	1.2	93.2
96.00	168	1.1	94.3
97.00	142	.9	95.2
98.00	156	1.0	96.3
99.00	122	.8	97.1
100.00	118	.8	97.8
101.00	125	.8	98.7
102.00	80	.5	99.2
103.00	72	.5	99.7
104.00	38	.3	99.9
105.00	9	.1	100.0

TABLE B-3

CUMULATIVE FREQUENCY DISTRIBUTION OF PROPOSED AFQT FOR THE 1984 IOT&E EXAMINEES TESTED WITH ASVAB FORMS 11,12B, AND 13

Value	Frequency	Percent	Cum Percent
7.00	1	.0	.0
10.00	1	.0	.0
12.00	5	.0	.0
13.00	1	.0	.0
14.00	2	.0	.0
15.00	8	.0	.0
16.00	10	.0	.0
17.00	20	.0	.1
18.00	23	.0	.1
19.00	40	.0	.1
20.00	45	.1	.2
21.00	63	.1	.3
22.00	87	.1	.4
23.00	111	.1	.5
24.00	115	.1	.6
25.00	166	.2	.8
26.00	182	.2	1.0
27.00	218	.3	1.3
28.00	232	.3	1.6
29.00	275	.3	1.9
30.00	320	.4	2.3
31.00	315	.4	2.6
32.00	323	.4	3.0
33.00	380	.4	3.5
34.00	411	.5	4.0
35.00	492	.6	4.5
36.00	490	.6	5.1
37.00	565	.7	5.8
38.00	598	.7	6.5
39.00	645	.8	7.2
40.00	668	.8	8.0
41.00	800	.9	9.0
42.00	877	1.0	10.0
43.00	862	1.0	11.0
44.00	924	1.1	12.1
45.00	957	1.1	13.2
46.00	981	1.2	14.4
47.00	1022	1.2	15.6
48.00	1124	1.3	16.9
49.00	1190	1.4	18.3
50.00	1210	1.4	19.8
51.00	1231	1.5	21.2
52.00	1313	1.5	22.8
53.00	1360	1.6	24.4
54.00	1275	1.5	25.9
55.00	1414	1.7	27.5

TABLE B-3 (Con't.)

Value	Frequency	Percent	Cum Percent
56.00	1492	1.8	29.3
57.00	1391	1.6	30.9
58.00	1483	1.7	32.7
59.00	1467	1.7	34.4
60.00	1520	1.8	36.2
61.00	1435	1.7	37.9
62.00	1492	1.8	39.6
63.00	1449	1.7	41.4
64.00	1536	1.8	43.2
65.00	1436	1.7	44.9
66.00	1533	1.8	46.7
67.00	1509	1.8	48.4
68.00	1385	1.6	50.1
69.00	1492	1.8	51.8
70.00	1447	1.7	53.5
71.00	1455	1.7	55.3
72.00	1404	1.7	56.9
73.00	1395	1.6	58.6
74.00	1437	1.7	60.2
75.00	1431	1.7	61.9
76.00	1357	1.6	63.5
77.00	1278	1.5	65.0
78.00	1407	1.7	66.7
79.00	1256	1.5	68.2
80.00	1292	1.5	69.7
81.00	1308	1.5	71.2
82.00	1296	1.5	72.8
83.00	1286	1.5	74.3
84.00	1203	1.4	75.7
85.00	1164	1.4	77.1
86.00	1273	1.5	78.6
87.00	1195	1.4	80.0
88.00	1204	1.4	81.4
89.00	1185	1.4	82.8
90.00	1145	1.3	84.1
91.00	1187	1.4	85.5
92.00	1111	1.3	86.9
93.00	1104	1.3	88.2
94.00	1087	1.3	89.4
95.00	1117	1.3	90.8
96.00	1049	1.2	92.0
97.00	974	1.1	93.1
98.00	1013	1.2	94.3
99.00	905	1.1	95.4
100.00	910	1.1	96.5
101.00	899	1.1	97.5
102.00	744	.9	98.4
103.00	651	.8	99.2
104.00	452	.5	99.7
105.00	245	.3	100.0

TABLE B-4

CONVERSION TABLE FOR SUM OF RAW SCORES
TO PERCENTILE SCORES ON THE PROPOSED AFQT

Sum of subtest raw scores	Percentile score	Sum of subtest raw scores	Percentile score
1-18	1	63	41
19-21	2	64	42
22-23	3	65	43
24	4	66	45
25	5	67	46
26	5	68	48
27	6	69	48
28	6	70	49
29	7	71	51
30	8	72	52
31	9	73	53
32	9	74	55
33	10	75	56
34	11	76	56
35	12	77	58
36	12	78	60
37	13	79	61
38	14	80	63
39	15	81	65
40	16	82	66
41	17	83	68
42	18	84	68
43	19	85	69
44	20	86	71
45	21	87	72
46	22	88	74
47	23	89	75
48	24	90	77
49	25	91	78
50	25	92	80
51	26	93	81
52	27	94	83
53	29	95	84
54	30	96	86
55	31	97	87
56	32	98	89
57	34	99	90
58	35	100	92
59	36	101	95
60	38	102	96
61	38	103	98
62	39	104	99
		105	99

APPENDIX C

ANALYSIS OF INITIAL-TEST EXAMINEES

APPENDIX C

ANALYSIS OF INITIAL-TEST EXAMINEES

As discussed in the main test, the ASVAB answer sheet has a section called "test type," in which examinees can mark if they are taking the test for the first time (INITIAL) or are being retested (RET or VERIFIC). Apparently many examinees during the 1984 IOT&E misunderstood the meaning of the response alternatives because many more of them marked that they were taking a retest than are ordinarily retested, according to the Military Enlistment Processing Command (MEPCOM) records. Because the responses made by the examinees could not be accurately interpreted, the main analysis included all applicants for active duty, those who said they were taking the test initially and those who said they were being retested. In this appendix, the sample of examinees is restricted to those who marked INITIAL on their answer sheets, people who marked other responses or omitted the section were deleted.

The amount of inflation was completed using the equipercentile equating technique explained in appendix A. The same weighting procedures were used for this sample as for the entire group (see appendix A).

The amount of inflation in the NO scores for this sample is shown in table C-1. The maximum inflation of the NO scores is 1 point lower for each sex in this group than reported in the main text. In general the amount of inflation of NO scores for this group is up to 1 point lower than for the entire group. Qualification rates are shown in tables C-2 through C-6 for the total DOD, Army, Navy, Air Force, and Marine Corps applicants, respectively. Because of the lower inflation of NO scores, rates for the adjusted AFQT scores are closer to those for the operational AFQT scores than was true for total samples of examinees.

TABLE C-1

EQUATING OF NO AND SCORES FOR EXAMINEES IDENTIFIED AS INITIAL TEST TAKERS

NO score in 1980 Youth Population	Equivalent score in 1984 IOT&E group	
	NO for males	NO for females
1	1-8	1-8
2	9	9
3	10	10
4	11	11
5	12	12
6	13	13
7	14	14
8	15	16
9	16	17
10	17	18
11	18	19
12	18	20
13	19	21
14	20	21
15	20	22
16	21	22
17	22	23
18	23	24
19	24	24
20	24	25
21	25	25
22	26	26
23	27	27
24	28	28
25	29	29
26	30	30
27	31	31
28	32	32
29	33	33
30	34	34
31	34	35
32	35	36
33	36	36
34	37	37
35	38	38
36	39	39
37	40	40
38	40	41
39	41	42
40	42	43

TABLE C-1 (Cont.)

NO score in 1980 Youth Population	Equivalent score in 1984 IOT&E group	
	NO for males	NO for females
41	43	44
42	44	45
43	44	46
44	45	47
45	46	47
46	47	47
47	47	48
48	48	48
49	49	49
50	50	50

TABLE C-2

QUALIFYING RATES ON ALTERNATIVE AFQTS
DEPARTMENT OF DEFENSE APPLICANTS

Group	Number of cases[a]	Percent qualifying on AFQT at percentile score: 21 Oper[b]	21 Adj[c]	21 Prop[d]	31 Oper[b]	31 Adj[c]	31 Prop[d]	50 Oper[b]	50 Adj[c]	50 Prop[d]
Males										
Total	43331	90.0	88.2	90.5	77.5	74.8	78.9	51.9	49.0	52.3
White	31128	94.8	93.6	95.4	85.9	83.5	87.6	61.9	58.7	62.1
Black	8152	76.3	72.6	76.7	53.0	49.2	54.5	22.6	20.4	23.2
Other	4051	80.5	77.8	81.2	62.4	59.8	65.3	33.9	31.6	35.3
Females										
Total	8110	92.4	90.8	91.6	80.3	77.6	78.5	52.5	49.9	48.3
White	5200	96.6	95.7	96.3	89.4	87.2	88.3	65.8	63.0	61.2
Black	2270	84.9	81.9	82.3	62.6	58.6	59.3	26.2	23.7	22.5
Other	640	85.2	82.8	86.4	69.4	67.0	66.9	38.4	36.3	35.6

---

a. Applicants for active duty tested October and November 1984 that marked 'INIT' as test type on their answer sheets.
b. Operational AFQT score, includes inflated NO scores.
c. Adjusted AFQT score, includes deflated NO scores.
d. Proposed AFQT score, with NO replaced by MK.

TABLE C-3

QUALIFYING RATES ON ALTERNATIVE AFQTS
ARMY APPLICANTS

Group	Number of cases[a]	Percent qualifying on AFQT at percentile score: 21 Oper[b]	21 Adj[c]	21 Prop[d]	31 Oper[b]	31 Adj[c]	31 Prop[d]	50 Oper[b]	50 Adj[c]	50 Prop[d]
Males										
Total	20252	86.4	84.1	87.1	72.0	69.0	73.8	45.9	43.1	46.5
White	13934	92.8	91.2	93.5	82.4	79.7	84.1	57.1	54.0	57.8
Black	4277	71.1	67.2	71.6	46.6	42.5	48.2	17.8	16.0	18.4
Other	2041	74.4	71.1	75.4	54.1	51.8	57.5	28.4	26.0	29.0
Females										
Total	3619	89.6	87.5	89.0	76.8	73.9	74.0	47.5	45.0	43.5
White	2087	95.3	93.9	95.2	88.0	85.8	86.4	63.4	60.4	59.0
Black	1224	82.0	78.9	80.0	60.5	56.4	55.9	23.9	21.9	20.4
Other	308	80.5	77.9	82.8	64.9	62.7	61.7	33.8	31.8	30.5

---

a. Applicants for active duty tested October and November 1984 that marked 'INIT' as test type on their answer sheets.
b. Operational AFQT score, includes inflated NO scores.
c. Adjusted AFQT score, includes deflated NO scores.
d. Proposed AFQT score, with NO replaced by MK.

TABLE C-4

QUALIFYING RATES ON ALTERNATIVE AFQTS
NAVY APPLICANTS

Group	Number of cases[a]	Percent qualifying on AFQT at percentile score: 21 Oper[b]	21 Adj[c]	21 Prop[d]	31 Oper[b]	31 Adj[c]	31 Prop[d]	50 Oper[b]	50 Adj[c]	50 Prop[d]
Males										
Total	10202	93.2	91.7	93.8	82.1	79.6	83.5	57.1	54.3	57.4
White	7942	96.2	95.2	96.8	88.0	85.6	89.0	64.7	61.8	64.9
Black	1412	81.7	77.8	81.6	58.5	55.2	60.5	26.4	23.7	26.4
Other	848	84.6	82.3	85.4	66.6	63.8	70.2	36.4	34.9	38.9
Females										
Total	1670	95.4	94.4	94.7	85.2	82.5	84.6	59.7	57.2	55.3
White	1213	97.9	97.2	97.4	91.2	89.3	91.0	70.3	67.9	65.5
Black	322	89.1	87.6	86.7	67.7	61.8	66.2	27.6	24.8	23.0
Other	135	88.2	85.9	89.6	72.6	71.1	71.1	40.7	38.5	40.7

---

a. Applicants for active duty tested October and November 1984 that marked 'INIT' for test type on their answer sheets.
b. Operational AFQT score, includes inflated NO scores.
c. Adjusted AFQT score, includes deflated NO scores.
d. Proposed AFQT score, with NO replaced by MK.

TABLE C-5

QUALIFYING RATES ON ALTERNATIVE AFQTS
AIR FORCE APPLICANTS

Group	Number of cases[a]	Percent qualifying on AFQT at percentile score: 21 Oper[b]	21 Adj[c]	21 Prop[d]	31 Oper[b]	31 Adj[c]	31 Prop[d]	50 Oper[b]	50 Adj[c]	50 Prop[d]
Males										
Total	7435	93.9	92.6	94.5	84.5	82.0	85.8	60.4	57.2	61.5
White	5521	97.1	96.5	97.5	90.6	88.5	91.3	68.7	65.3	69.2
Black	1324	83.0	79.5	84.4	63.8	59.8	66.1	32.7	29.8	34.7
Other	590	87.8	85.9	88.8	74.6	71.5	79.0	44.8	42.4	49.3
Females										
Total	2325	94.0	92.6	92.9	81.2	78.8	79.9	53.4	50.5	49.0
White	1581	97.0	96.5	96.5	89.1	86.7	88.0	63.9	61.1	59.1
Black	587	86.5	83.3	83.5	62.0	59.5	60.3	27.8	24.7	24.5
Other	157	91.1	88.5	91.1	72.6	70.7	71.3	44.0	40.8	39.5

---

a. Applicants for active duty tested October and November 1984 who marked 'INIT' as test type on their answer sheets.
b. Operational AFQT score, includes inflated NO scores.
c. Adjusted AFQT score, includes deflated NO scores.
d. Proposed AFQT score, with NO replaced by MK.

TABLE C-6

QUALIFYING RATES ON ALTERNATIVE AFQTS
MARINE CORPS APPLICANTS

Group	Number of cases[a]	Percent qualifying on AFQT at percentile score: 21 Oper[b]	21 Adj[c]	21 Prop[d]	31 Oper[b]	31 Adj[c]	31 Prop[d]	50 Oper[b]	50 Adj[c]	50 Prop[d]
Males										
Total	4943	92.1	90.6	92.2	79.8	77.6	79.6	52.0	49.1	50.6
White	3330	95.8	94.9	96.1	87.6	85.7	87.7	63.0	59.8	61.1
Black	1093	81.4	78.3	81.2	58.5	55.4	57.8	24.3	22.6	24.3
Other	520	90.6	89.2	90.0	74.6	72.3	73.9	40.0	36.2	38.7
Females										
Total	430	96.1	94.4	94.7	86.3	83.3	83.5	59.5	57.0	55.6
White	267	98.1	97.0	98.9	92.3	91.8	91.0	72.7	70.8	68.5
Black	128	93.0	89.8	86.7	71.1	65.6	68.0	34.4	29.7	30.5
Other	35	91.4	91.4	91.4	88.6	82.9	82.9	51.4	51.4	48.6

---

a. Applicants for active duty tested October and November 1984 who marked 'INIT' as test type on their answer sheets.
b. Operational AFQT score, includes inflated NO scores.
c. Adjusted AFQT score, includes deflated NO scores.
d. Proposed AFQT score, with NO replaced by MK.